S. HRG. 108–359

CLIMATE HISTORY AND THE SCIENCE UNDERLYING FATE, TRANSPORT, AND HEALTH EFFECTS OF MERCURY EMISSIONS

HEARING

BEFORE THE

COMMITTEE ON ENVIRONMENT AND PUBLIC WORKS UNITED STATES SENATE

ONE HUNDRED EIGHTH CONGRESS

FIRST SESSION

JULY 29, 2003

Printed for the use of the Committee on Environment and Public Works

U.S. GOVERNMENT PRINTING OFFICE

92–381 PDF WASHINGTON : 2004

For sale by the Superintendent of Documents, U.S. Government Printing Office
Internet: bookstore.gpo.gov Phone: toll free (866) 512–1800; DC area (202) 512–1800
Fax: (202) 512–2250 Mail: Stop SSOP, Washington, DC 20402–0001

COMMITTEE ON ENVIRONMENT AND PUBLIC WORKS

ONE HUNDRED EIGHTH CONGRESS
FIRST SESSION

JAMES M. INHOFE, Oklahoma, *Chairman*

JOHN W. WARNER, Virginia
CHRISTOPHER S. BOND, Missouri
GEORGE V. VOINOVICH, Ohio
MICHAEL D. CRAPO, Idaho
LINCOLN CHAFEE, Rhode Island
JOHN CORNYN, Texas
LISA MURKOWSKI, Alaska
CRAIG THOMAS, Wyoming
WAYNE ALLARD, Colorado

JAMES M. JEFFORDS, Vermont
MAX BAUCUS, Montana
HARRY REID, Nevada
BOB GRAHAM, Florida
JOSEPH I. LIEBERMAN, Connecticut
BARBARA BOXER, California
RON WYDEN, Oregon
THOMAS R. CARPER, Delaware
HILLARY RODHAM CLINTON, New York

ANDREW WHEELER, *Majority Staff Director*
KEN CONNOLLY, *Minority Staff Director*

C O N T E N T S

Page

JULY 29, 2003

OPENING STATEMENTS

Allard, Hon. Wayne, U.S. Senator from the State of Colorado, prepared statement 11
Cornyn, Hon. Jon, U.S. Senator from the State of Texas, prepared statement . 58
Inhofe, Hon. James M., U.S. Senator from the State of Oklahoma 1
Jeffords, Hon. James M., U.S. Senator from the State of Vermont, prepared statement 7
Voinovich, Hon. George V., U.S. Senator from the State of Ohio 3

WITNESSES

Legates, David R., director, Center for Climatic Research, University of Delaware 12
Prepared statement 209
Levin, Leonard, program manager, Electric Power Research Institute 40
Prepared statement 211
Mann, Michael E., assistant professor, University of Virginia, Department of Environmental Sciences 9
Prepared statement 173
Responses to additional questions from:
Senator Inhofe 178
Senator Jeffords 194
Myers, Gary, professor of Neurology and Pediatrics, Department of Neurology, University of Rochester Medical Center 44
Prepared statement 299
Rice, Deborah C., toxicologist, Bureau of Remediation and Waste Management, Maine Department of Environmental Protection 42
Prepared statement 283
Responses to additional questions from Senator Jeffords 284
Soon, Willie, astrophysicist, Harvard-Smithsonian Center for Astrophysics 6
Prepared statement 58
Responses to additional questions from Senator Jeffords 155

ADDITIONAL MATERIAL

Articles:
Climate Research, Vol. 23:89–110, 2003, Proxy Climatic and Environmental Changes of the Past 1000 years 127–148
Energy & Environment Vol. 14, Nos. 2 and 3, 2003, Reconstructing Climatic and Environmental Changes of the Past 1000 Years: A Reappraisal 60–126
Geophysical Research Letters, Vol. 31, Estimation and Representation of Long-term (>40 year) Trends of Northern-Hemisphere-gridded Surface Temperature: A Note of Caution 149–154
Original Contributions, Effects of Prenatal and Postnatal Methylmercury Exposure From Fish Consumption on Neurodevelopment 302–308
Personal Health, Tip the Scale in Favor of Fish: The Healthful Benefits Await 321
Risk Analysis, Vol. 23, No. 1, 2003, Methods and Rationale for Derivation of a Reference Dose for Methylmercury by the U.S. EPA 290–298

Page

—Continued
The Atlanta Journal-Constitution, June 6, 2003, Clear Skies Mercury Curb Put in Doubt 317
The Lancet, Prenatal Methylmercury Exposure from Ocean Fish Consumption in the Seychelles Child Development Study.......... 309–316
The New York Times, July 29, 2003, Does Mercury Matter? Experts Debate the Big Fish Question 319
The Philadelphia Inquirer, March 7, 2003, Mercury Rising 318
Chart, National Mean Mercury Concentration in Tissues of Selected Fish Species (all sample types) 289
Letter, to Senator Inhofe, from John Christy 323
Report, EPRI, May 2003, A Framework for Assessing the Cost-Effectiveness of Electric Power Sector Mercury Control Policies 217–282

CLIMATE HISTORY AND THE SCIENCE UNDERLYING FATE, TRANSPORT, AND HEALTH EFFECTS OF MERCURY EMISSIONS

TUESDAY, JULY 29, 2003

U.S. SENATE,
COMMITTEE ON ENVIRONMENT AND PUBLIC WORKS,
Washington, DC.

The committee met, pursuant to notice, at 9 o'clock a.m. in room 406, Senate Dirksen Building, Hon. James M. Inhofe (chairman of the committee) presiding.

Present: Senators Inhofe, Allard, Carper, Clinton, Cornyn, Jeffords, Thomas and Voinovich.

OPENING STATEMENT OF HON. JAMES M. INHOFE, U.S. SENATOR FROM THE STATE OF OKLAHOMA

Senator INHOFE. The meeting will come to order.

We have a policy that we announced when I became chairman of the committee that we will start on time, whether anyone is here or not here, members, witnesses or others. So I appreciate all of you being punctual in spite of the fact that the Senators are not.

One of my primary objectives as chairman of the committee is to improve the way in which science is used. I think that when I became chairman of this committee, I announced three very outrageous things that we were going to do in this committee that have not been done before. No. 1, we are going to try to base our decisions, things that we do, on sound science. No. 2, we are going to be looking at the costs of some of these regulations, some of these policies that we have, and determine what they are going to be. And No. 3, we are going to try to reprogram the attitudes of the bureaucracy so that they are here not to rule, but to serve.

Good public policy decisions depend on what is real or probable, not simply on what serves our respective political agendas. When science is debated openly and honestly, public policy can be debated on firmer grounds. Scientific inquiry cannot be censored. Scientific debate must be open. It must be unbiased. It must stress facts rather than political agendas.

Before us today, we have two researchers who have published what I consider to be a credible, well-documented, and scientifically defensible study examining the history of climate change. Furthermore, these are top fields of inquiry in the Nation's energy environment debate and really the entire world's energy environment debate. We can all agree that the implications of this science are global, not only in terms of the environmental impacts, but also en-

ergy impacts, global trade impacts, and quite frankly, no less than global governance impacts.

We could also all agree that as a result of the import and impact of these issues, it is absolutely crucial that we get this science right. False or incomplete or misconstrued data are simply not an acceptable basis for policymaking decisions in which the Congress of the United States is involved. Such data would violate the Data Quality Act, which we passed on a bipartisan basis here in the Senate and which we have bipartisanly embraced. If we need more data to satisfy our standards, then so be it.

This Administration is prepared to do so in an aggressive strategy that the climate change strategic plan outlines. The 1000-year climate study that the Harvard-Smithsonian Center for Astrophysics has compiled is a powerful new work of science. It has received much attention, and rightfully so. I would add at this time, it did not receive much attention from some of the liberal media who just did not want to believe that any of the facts that were disclosed were accurate.

I think the same can be said in terms of work that has recently received attention of the hockey stick study. In many important ways, the Harvard-Smithsonian Center's work shifts the paradigm away from the previous hockey stick study. The powerful new findings of this most comprehensive study shiver the timbers of the adrift Chicken Little crowd.

I look forward to determining whose data is most comprehensive, uses the most proxies, maintains the regional effects, avoids losing specificity through averaging statistics, considers more studies, and most accurately reflects the realities of the Little Ice Age, reflects the realities of the Medieval Warming Period, and more.

Mercury presents a different set of issues. That would be our second panel. It is well-established that high levels of exposure to methyl-mercury before birth can lead to neuro-development problems. But what about mercury consumed through fish, the most common form of prenatal exposure? Mercury makes its way into fish through various ways, but primarily though deposition from air emissions, with 80 percent of emissions deposited either regionally or globally, not locally. Global mercury emissions are about 5,000 tons a year. About half of those are man-made emissions.

In the United States, a little more than 100 tons are emitted from non-power plant sources. Industry is making great strides in reducing these emissions. I would like to submit for the record this EPA document available on their Web site which indicates that when rules now on the books are fully implemented at non-power plant, nationwide emissions will be cut by nearly 50 percent. Power plants emit about 50 tons of mercury annually, about 1 percent of the worldwide emissions.

In setting policy, key questions need to be answered, such as how would controls change this deposition; what portion of mercury exposure can not be controlled; and what are the health impacts of prenatal exposure. We will hear testimony today that indicates any changes to mercury exposure in fish would be minimal under even the most stringent proposal to regulate mercury. Today, we will also hear testimony that the most recent and comprehensive study

to date found no evidence that prenatal mercury exposure from ocean fish presents a neurological risk.

So we have diverse opinions that will be discussed today, and that is the reason for this hearing, to wade through that so that those on the panel that will be making policy decisions will understand. I think it is no secret that we are not scientists up here, so we look at things logically.

With that, I would recognize one of my colleagues here that I have a great deal of respect for. Senator Voinovich and I started out together as we were mayors of cities almost 25 years ago. I consider him to be one of the real experts in the area of air. In fact, I can remember calling him in as an expert when he was Governor of Ohio and we were holding these hearings and I was chairman at that time of the Clean Air Subcommittee. I would recognize Senator Voinovich for any comments he would like to make or opening statements.

OPENING STATEMENT OF HON. GEORGE V. VOINOVICH, U.S. SENATOR FROM THE STATE OF OHIO

Senator VOINOVICH. Thank you, Mr. Chairman.

I want to congratulate you for the very comprehensive floor speech that you gave yesterday on the issue of climate change.

Senator INHOFE. I guess I should apologize. It was 12,000 words and I know you were anxious to get some floor time, so I appreciate your patience.

Senator VOINOVICH. Your words were much more scientifically based than mine.

[Laughter.]

Senator VOINOVICH. The two issues that we are going to explore at the hearing today, the science of mercury and the science of climate change, are both important and timely. I commend you for holding this hearing.

I think I do not have to remind you that we have had hearings on climate change now during the last 4 or 5 years. I think I had a couple when I was chairman of even the Transportation Infrastructure Committee. Senator Lieberman had hearings over in Governmental Affairs when he was chairman of the committee a year or so ago. So it is not a subject that is brand new to this committee.

I have stated time and time again here in the committee and on the floor that we must recognize that energy policy and environmental policy are two sides of the same coin, and the Senate has responsibility to harmonize these policies. We have an obligation here in the committee to ensure that legislation that we consider will protect our environment. We also have an obligation to ensure that any legislation we consider takes into account its potential impact on our economy and we have a moral obligation to ensure that we consider a bill's particular impact on the poor and the elderly who must survive on fixed incomes.

When the Senate takes up consideration of climate change and multi-pollutant legislation, we must keep that moral obligation in mind. We must ensure that we do not pass legislation that will significantly drive up the cost of electricity and home heating for those who can least afford them.

Several members of this committee have introduced pieces of legislation this year to reduce power plant emissions, including mercury, and address the issue of carbon emissions and climate change by capping carbon. Examples include Jeffords-Lieberman four-P bill, the Carper four-P bill, and the McCain-Lieberman climate change bill, which I understand will likely be offered as an amendment to the energy bill, just this week we are going to be considering it.

These bills will establish a nationwide cap on carbon emissions and their passage would force the utility sector, that is now using coal to generate over half of our Nation's electricity. To rely solely on natural gas for generation, we will have fuel switching—capping carbon equals fuel switching equals no-coal—to rely on natural gas regeneration despite the fact we have over a 250-year supply of domestic coal and are currently in the grips of a natural gas crisis in this country.

This crisis is a result of environmental policies that have driven up the use of natural gas in electricity generation significantly, while domestic supplies of natural gas have fallen, partly because we cannot do the exploration that we need to do for natural gas.

The result is predictable: tightening supplies of natural gas, higher natural gas prices, and higher electricity prices. Home heating prices are up dramatically, forcing folks on low and fixed incomes to choose between heating their homes and paying for other necessities such as food or medicine. The language that has been offered by Senators Jeffords, McCain, Lieberman and Carper if enacted will force our utilities to fuel switch to natural gas; will significantly raise energy prices; and will cause thousands of jobs to be lost, particularly in manufacturing States like my State of Ohio, which is already under duress in terms of manufacturing.

During the debate last year on the Jeffords-Lieberman four-P bill, I put together a white paper that discussed the impact that the bill would have if it were enacted. The numbers are staggering: an overall reduction in GDP of $150 billion by 2020, the loss of over 900,000 jobs by 2020, and a decline in national household earnings of $550 annually.

The cost of climate-change language such as the McCain-Lieberman bill could come without any benefits to our air quality or public health. Not even the most ardent supporter, and I hope this comes up, of carbon regulation will claim that there are demonstrable health benefits from carbon regulation. Yet the Energy Information Administration estimates that the passage of the McCain-Lieberman bill, if enacted, will raise petroleum product prices by 31 percent, raise natural gas prices by 79 percent, raise electricity prices by 46 percent, and reduce GDP by up to $93 billion by 2025.

Carbon caps and unrealistic mercury caps means fuel switching, again. The fuel switching means the end of manufacturing in my State, enormous burdens on the least of our brethren. It means moving jobs and production overseas, where there are less stringent environmental programs. And will actually, if you really think about it, increase global levels of pollution.

The question we face in this committee is whether we should do something reasonable to improve our understanding of the issues

surrounding carbon emissions and climate change, and attempt to reduce atmospheric concentration of carbon and mercury emissions without harming our economy, or rush into short-sighted policy that will cap carbon and mercury at unreasonable levels, shut down our economy, cut thousands of jobs, and move manufacturing overseas.

In a recent column, former Secretary of Energy James Schlesinger commented that:

> "In climate change, we have only a limited grasp of the overall forces at work. Uncertainties have continued to abound and must be reduced. In any approach to policy formation, this is very important, under conditions of such uncertainty should be taken only on an exploratory or a sequential basis. A premature commitment to a fixed policy could only proceed with fear and trembling."

I would like to have that column inserted in the record, Mr. Chairman.

Senator INHOFE. Without objection, so ordered.

Senator VOINOVICH. As I mentioned previously once or twice, I am working with Chairman Inhofe and the Administration on moving Clear Skies forward, which I intend to mark up in my subcommittee this fall. I am currently working with business and environmental groups to find a bipartisan compromise on dealing with carbon and global warming, with an emphasis on sound science, carbon sequestration, development of clean coal technologies, and a responsible approach that focuses more on consensus rather than politics.

We need more Senators to focus on moving forward in a responsible way and move away from harshly ideological positions that advance nothing other than the agenda of some environmental groups that have made carbon cap a political litmus test.

I thank the chairman for holding this important hearing and I look forward to hearing the testimony from our witnesses.

Senator INHOFE. That is an excellent opening statement, Senator Voinovich. I go back to one of your first sentences when you talked about the number of hearings we have had. We have to keep in mind that each new hearing has new data. For example, the 1,000-year Harvard-Smithsonian was not even out until March of this year. So there are new things that are coming along and I see a new trend-line which I discussed on the House of the Senate yesterday. So this will be a very valuable hearing.

Senator Cornyn, would you have any opening statement to make?

Senator CORNYN. I would like to reserve any statement until later, Mr. Chairman.

Senator INHOFE. Yes, that is fine. First, I would like to ask the first panel to come up. Dr. Legates, Dr. Willie Soon and Dr. Mann, would you three come up? First of all, we are honored to have who I consider three very excellent and professional scientific witnesses here today. Normally, we restrict the opening statements to 5 minutes, but it would be fine if you want to go about 7 minutes because I know you have come a long way and what we are dealing with here is probably one of the most significant things facing America, facing our economy, facing our environment today.

So I would introduce all three. Dr. David Legates is the director of the Center for Climatic Research at the University of Delaware.

Dr. Willie Soon is the astrophysicist at Harvard-Smithsonian Center for Astrophysics, and Dr. Michael Mann is assistant professor at the University of Virginia Department of Environmental Sciences. I will first ask Dr. Willie Soon to give his opening statement.

STATEMENT OF WILLIE SOON, ASTROPHYSICIST, HARVARD-SMITHSONIAN CENTER FOR ASTROPHYSICS

Dr. SOON. Mr. Chairman, distinguished Senators, my fellow panelists, Dr. Mann and Dr. Legates, and members of the audience, my name is Willie Soon. About a month or two ago, I became a very proud and grateful U.S. citizen. I just cannot believe where I am sitting today.

I am an astrophysicist with the Harvard-Smithsonian Center for Astrophysics in Cambridge, Massachusetts. My training is in atmospherics and space physics. My research interests for the past 10 years include changes in the sun and their possible impact on climate.

I am here today to testify that the climate of the 20th century is neither unusual nor the most extreme. Around 1,000 years ago, the temperature over many parts of the world was warm. A widespread cooling then set in for several centuries, followed by a recovery to 20th century warming.

My colleague and I collected the information on climate by proxy. We studied environmental indicators of local climate change going back some 1,000 years from many locations around the world. Based on work of approximately 1,000 researchers and hundreds of peer-reviewed papers, we conclude the following three points about climate history of the last 1,000 years.

On a location-by-location basis, point No. 1, there was warming from 800 to 1300 A.D., all about 1,000 years ago, over many parts of the world. This period is called the Medieval Warm Period. Following the warming of 1,000 years ago was a general cooling from about 1300 to 1900 A.D. This period is called the Little Ice Age.

Point No. 2, there is no convincing evidence from local proxy to suggest that the 20th century had higher temperatures or more extreme climate than the warm period 1,000 years ago.

Point No. 3, local and regional, rather than global average changes are the most relevant and practical measure of climate changes and its impact. Much of the climate proxy results using our work are new. Most papers were published in the scientific literature in the recent 5 to 10 years. There are two points to note about our methods. First, we keep the local or regional information contained in each climate proxy. This is important for studying geographical patterns of climate, which does not change everywhere at the same time.

Second, climate is more than just temperature, so we keep the climate information like rainfall, expansion or contraction of forests, all advances or retreats of glaciers, et cetera. Our approach makes use of the richness of information in climate proxies, which map out local environmental and climate properties, rather than just temperature alone.

The entirety of climate proxies over the last 1,000 years shows that over many areas of the world, there has been and continues

to be large local climatic changes. Those changes provide important changes for the computer simulations of climate. The full models which explore the Earth region by region can be tested against the natural patterns of change over the last 1,000 years that are detailed by the climate proxies.

Having computer simulation, we produced past patterns of climate which has been influenced predominantly by natural factors and is key to making an accurate forecast that includes all potential human-made warming and cooling effects.

In summary, based on expert conclusions from climate proxies in several hundred peer-reviewed papers by over 1,000 researchers from around the world, we find the following. No. 1, from one location to another, large natural swings in climate have occurred over the last 1,000 years. Those patterns have not always been synchronous.

No. 2, there was widespread warmth about 1,000 years ago, followed by widespread cooling ending by the beginning of the 20th century.

No. 3, the local and regional climate proxies cannot confirm that the 20th century is the warmest or most extreme over much of the world, compared especially to the Medieval Warm Period approximately 1,000 years ago.

This is all for my oral remarks and I thank you for the opportunity to be here.

Senator INHOFE. Dr. Soon, we appreciate that excellent opening statement. You did not even take all of your time. That is very unusual.

At this time, Dr. Mann if you don't mind, I would like to interrupt your testimony. We have been joined by the Ranking Minority Member, Senator Jeffords. Senator Jeffords, do you have an opening statement you would like to make at this time?

Senator JEFFORDS. I would ask unanimous consent that it be made as part of the record and would prefer listening to the witnesses.

[The prepared statement of Senator Jeffords follows:]

STATEMENT OF HON. JAMES JEFFORDS, U.S. SENATOR FROM THE STATE OF VERMONT

We're here today to discuss two very important topics—climate change and mercury pollution. As most of you know, I am the author of ambitious legislation—the Clean Power Act of 2003—which addresses these environmental problems, as well as ozone, acid rain, and human health damage from fine particulate matter.

Unfortunately, we aren't here today to talk about moving forward to find innovative solutions to these real world problems. Instead, today's hearing will largely be a mirror or the reverse of the robust and growing consensus in the mainstream scientific community on climate and mercury pollution.

The disappointing result will be more delay. Delay on the part of Congress, and even worse, the ongoing backsliding on the part of the Administration, means that we fail to act responsibly as a society to protect future generations. That means increasingly greater risks of global warming and mercury poisoning.

There is no doubt that the scientific process must inform policymakers as new information comes in. Unfortunately, there is no new information to be found here today that would dissuade us from acting quickly and responsibly to reduce greenhouse gas and mercury emissions. In today's discussion of a literature survey of climate research, the skeptics are trotting out an argument that is several years old and already discarded by their peers.

It is abundantly clear that now is the time to act.

• The National Academy of Sciences has said, "Despite the uncertainties, there is general agreement that the observed warming is real and particularly strong within the past 20 years."

• NOAA currently says that,

> "The climatic record over the last thousand years clearly shows that global temperatures increased significantly in the 20th Century, and that this warming was likely to have been unprecedented in the last 1200 years."

• EPA's website says that, "There is new and stronger evidence that most of the warming over the last 50 years is attributable to human activities."

One would have to be madder than a March hare to fail to see the need to act. Yet, the Administration's new research plan falls squarely into hare territory—denying the reality staring them in the face.

I want to show you the latest odds on warming. MIT says that there is a one in five chance that the temperature of the earth will warm by approximately 4 or 5 degrees over the course of this century, assuming there is no action to reduce emissions.

As my dear departed friend, Senator John Chafee, said in 1989:

> "It is clear that we are facing a serious threat. The scientists are telling us that if we continue to stroll along as if everything is fine, we will transform Earth into a planet that will not be able to support life as we now know it."

While mercury contamination does not have the same dramatic effect on earth's systems, it is still a dangerous global and local pollutant because it is bio-accumulative and toxic to human health.

Long ago, Congress decided that toxic air emissions should be reduced and took very aggressive steps in 1990 to make that happen, especially if they fall into the Great Lakes and other great waters like Lake Champlain. Unfortunately, the Agency has fallen significantly behind in complying with the Clean Air Act's schedule. A settlement agreement mandates controlling toxic air pollutants from utilities by 2008.

In 1998, related to the controversy around EPA's late reports to Congress on utility air toxics, Congress directed the National Academy of Sciences (NAS) to recommend an appropriate reference dose for mercury exposure. In 2000, the NAS reported that EPA's reference dose was scientifically sound and adequate to protect most Americans. That NAS review considered all health effects studies, including the Seychelles study that we'll discuss today.

We know that mercury is a potent toxic. It affects the human brain, spinal cord, kidneys, liver and the heart. It affects the ability to feel, see, taste and move. We know that mercury can affect fetal development, preventing the brain and nervous system from developing normally. Long term exposure to mercury can result in stupor, coma and personality changes.

"Mad as a Hatter" is the phrase that was used in the 1800's to describe the employees of the felt hat industry whose constant exposure to mercury changed their behavior. Fortunately, Americans exposure from commercial and recreational fish consumption is substantially less than that, though dozens of health warnings are posted nationwide.

But, it's crazy for anyone to suggest that we should not reduce mercury emissions significantly, since we know its health effects and we have the technologies to control it.

We should have a hearing on how to export those control technologies and Congress should urge the Administration to negotiate binding global reductions in mercury, as the Senate did last year in the Energy bill for greenhouse gas emissions.

At a minimum, we should pass four-pollutant legislation now that gets reductions faster and deeper than required by the current Clean Air Act. I'm sad to say that there have been no negotiations on that front since I initiated some in early 2002. And the Administration has done nothing to reduce these emissions with its abundant authority in the Act.

We can't afford to leave these problems to future generations to solve. We can't let our children and grandchildren wake up to find that our delays have cost them dearly in terms of health and the global and local environment. It's time to act responsibly.

Finally, I ask that material from the journal EOS, the NOAA website, the Atlanta Journal Constitution, the National Center for Atmospheric Research, and the American Geophysical Union be included in the hearing record.

Senator JEFFORDS. I might point out, we have got to do something about this traffic out there.

[Laughter.]

Senator INHOFE. Well, the name of our subcommittee is Transportation and Infrastructure, so maybe we can do something about the traffic out there.

Senator JEFFORDS. I hope so.

Senator INHOFE. Dr. Mann, you are recognized.

STATEMENT OF MICHAEL E. MANN, ASSISTANT PROFESSOR, UNIVERSITY OF VIRGINIA, DEPARTMENT OF ENVIRONMENTAL SCIENCES

Dr. MANN. Senators, my name is Michael Mann. I am a professor in the Department of Environmental Sciences at the University of Virginia. My research involves the study of climate variability and its causes. I was a lead author of the IPCC Third Scientific Assessment report. I am current organizing committee chair for the National Academy of Sciences' Frontiers of Science, and have served as a committee member or adviser for other National Academy of Sciences' panels.

I have served as editor for the *Journal of Climate* of the American Meteorological Society for 3 years and I am a member of the advisory panel for the NOAA Climate Change Data and Detection Program. I am a member of numerous other international and U.S. scientific working groups, panels and steering committees. I have coauthored more than 60 peer-reviewed publications on diverse topics within the fields of climatology and paleoclimatology.

Honors I have received include selection in 2002 as one of the 50 leading visionaries in science and technology by Scientific American magazine, and the outstanding scientific publication award of NOAA for 2000.

In my testimony here today, I will explain, No. 1, how mainstream climate researchers have come to the conclusion that late 20th century warmth is unprecedented in a very long-term context and that this warmth is likely related to the activity of human beings; and No. 2, why a pair of recent articles challenging these conclusions by astronomer Willie Soon and his coauthors are fundamentally unsound.

It is the consensus of the climate research community that the anomalous warmth of the late 20th century cannot be explained by natural factors, but instead indicates significant anthropogenic, that is human influences. This conclusion is embraced by the position statement on climate change and greenhouse gases of the American Geophysical Union, by the 2001 report of the IPCC, the Intergovernmental Panel on Climate Change, and by a National Academy of Sciences' report that was solicited by the Bush Administration in 2001.

More than a dozen independent research groups have now reconstructed the average temperature of the northern hemisphere in past centuries, both by employing natural archives of past climate information or proxy indicators such as tree rings, corals, ice cores, lake sediments and historical documents, and through the use of climate model simulations. If I can have the first exhibit here, as shown in this exhibit, the various proxy reconstructions agree with

each other, as well as with the model simulations, all of which are shown, within the estimated uncertainties. That is the gray-shaded region.

The proxy reconstructions, taking into account these uncertainties, indicate that the warming of the northern hemisphere during the late 20th century, that is the northern hemisphere, not the globe, as I have sometimes heard my study incorrectly referred to, the northern hemisphere during the late 20th century, that is the end of the red curve, is unprecedented over at least the past millennium and it now appears based on peer-reviewed research, probably the past two millennia.

The model simulations demonstrate that it is not possible to explain the anomalous late-20th century warmth without the contribution from anthropogenic influences. These are the consensus conclusions of the legitimate community of climate and paleoclimate researchers investigating such issues.

Astronomers Soon and Baliunas have attempted to challenge the scientific consensus based on two recent papers, henceforth collectively referred to as SB, that completely misrepresent the past work of other legitimate climate researchers and are deeply flawed for the following reasons. No. 1, SB make the fundamental error of citing evidence of either wet or dry conditions as being in support of an exceptional Medieval Warm Period. Such an ill-defined criterion could be used to define any period of climate as either warm or cold. It is pure nonsense.

Experienced paleoclimate researchers know that they must first establish the existence of a temperature signal in a proxy record before using it to try to reconstruct past temperature patterns. If I can have exhibit two, this exhibit shows a map of the locations of a set of records over the globe that have been rigorously analyzed by my colleagues and I for their reliability as long-term temperature indicators. I will refer back to that graphic shortly.

No. 2, it is essential to distinguish between regional temperature changes and truly hemispheric or global changes. Average global or hemispheric temperature variations tend to be far smaller in their magnitude than those for particular regions. This is due to a tendency for the cancellation of simultaneous warm and cold conditions in different regions, something that anybody who follows the weather is familiar with, in fact.

As shown by exhibit three, if I can have that up here as well now, thank you, this exhibit plots the estimated temperature for various locations shown in the previously displayed map. As you can see, the specific periods of relative cold and warm, blue and red, differ greatly from region to region. Climatologists, of course, know this. What makes the late 20th century unique is the simultaneous warmth indicated by nearly all the long-term records. It is this simultaneous warmth that leads to the anomalous late-20th century warmth evident for northern hemisphere average temperatures.

The approach taken by SB does not take into account whether warming or cooling in different regions is actually coincident, despite what they might try to tell you here today.

No. 3, as it is only the past few decades during which northern hemisphere temperatures have exceeded the bounds of natural var-

iability, any analysis such as SB that compares past temperatures only to early or mid-20th century conditions; you repeatedly hear Dr. Soon refer to the 20th century; climatologists do not consider that a meaningful baseline because there has been a dramatic warming during the 20th century and the early 20th century and the late 20th century are almost as different as the late 20th century and any other period during the past 1,000 years at least. So a study that refers only to early or mid-20th century conditions or generic 20th century conditions and does not specifically address the late 20th century, cannot address the issue of whether or not late-20th century warmth is anomalous in a long-term context.

To summarize, late-20th century warming is unprecedented in modern climate history at hemispheric scales. A flawed recent claim to the contrary by scientists lacking expertise in paleoclimatology is not taken seriously by the scientific community.

The anomalous recent warmth is almost certainly associated with human activity and this is the robust consensus view of the legitimate climate research community.

Thank you.

Senator INHOFE. Thank you, Dr. Mann.

Dr. Legates.

First, I would ask Senator Allard, did you want to make an opening statement?

Senator ALLARD. Mr. Chairman, I do have an opening statement and in deference to the panel and you I would just like to have it put in the record. If you would do that, then I would be happy.

Senator INHOFE. Without objection.

[The prepared statement of Senator Allard follows:]

STATEMENT OF HON. WAYNE ALLARD, U.S. SENATOR FROM THE STATE OF COLORADO

Mr. Chairman, I want to thank you for holding this important hearing today.

As a veterinarian, I have some scientific training in my background. I strongly believe that we should use scientific principals as a guidepost when formulating any regulation. This scientific guidepost approach is particularly important when looking at regulations with the implications and magnitude of regulations on climate change and mercury control.

Climate change has been an ongoing discussion for many years. However, during the 1970's the concerns were exactly opposite of what they are now. Then we were told that there was a threat of massive global cooling. Headlines screamed that we were in danger of entering another ice age. Now we are told that massive warming trends are going to cause overheating across the globe. We need answers, not rhetoric.

All of the witnesses here today have a great deal of experience. All of the witnesses here have spent many years analyzing data related to the areas of their expertise. But, I am concerned that, at times, data may be reviewed selectively and in isolation. I am also concerned that emphasis may fall on a limited number of studies. In science we have all learned that the only way to solidly prove a theory is by conducting tests, studies or experiments that repeatedly arrive at the same result. We cannot simply ignore the studies that do not have the outcome we are looking for. This applies whether we are looking at climate change, mercury or any other issue.

I want to spend most of my time and attention today on potential mercury regulations. While today's hearing is intended to focus on science, I would also like to touch on the impact that potential regulations will have on the economy of my state and the west. As many of you know, western coal differs from other types of coal in several ways. The higher chlorine content in western coal makes it more difficult to remove mercury when burning it. And, while western coal does contain mercury, when it is burned it gives off mercury in the elemental form. It is my understanding that this is not the type of mercury that deposits in the ecosystem to potentially be absorbed by the environment.

The economies of Colorado, and the entire west, will be impacted by harsh regulations placed on their coal. Economies undoubtedly will be damaged by the decrease in use of coal mined in the West. In addition, while jobs are being lost due to the subsequent inability to fully utilize western coal supplies, if power can no longer be generated by using coal mined in the west, other less efficient coal types will have to be transported across long distances. This additional expenditure will add to the price of electricity generation, driving up electricity costs and further damaging an economy that will already be struggling.

This is why it is so important to me that we be cautious when dealing with situations such as these and why we should place strong emphasis on the use of sound science. Our regulations must be thoughtful reflections of what we know—they should not be reflexive or reactive attempts to legislate a cure before we know what the disease is.

Again, Mr. Chairman, thank you for holding this hearing. I look forward to hearing the witness testimony and discussions to come.

Senator INHOFE. That being the case, let's dispense with any further opening statements.

Dr. Legates, thank you very much for being here. You are recognized.

STATEMENT OF DAVID R. LEGATES, DIRECTOR, CENTER FOR CLIMATIC RESEARCH, UNIVERSITY OF DELAWARE

Dr. LEGATES. Thank you. Mr. Chairman, Distinguished Senators, Doctors Mann and Soon, and members of the audience, I would like to thank the committee for inviting my commentary on this important topic of climate history and its implications. My research interests have focused on hydroclimatology. That is the study of water in the atmosphere and on the land, and as well as on the application of statistical methodology in climatological research.

I am familiar with the testimony presented here by Dr. Soon. My contributions to Dr. Soon's research stem from my grappling with the striking disagreement between the longstanding historical record and the time series recently presented by Dr. Mann and his colleagues. It also stems from my own experiences in compiling and merging global estimates of air temperature and precipitation from a variety of disparate sources.

My Ph.D. dissertation resulted in the compilation of high-resolution climatologies of global air temperature and precipitation. From that experience, I have become acutely aware of the issues associated with merging data from a variety of sources and containing various biases and uncertainties. By its very nature, climatological data exhibit a number of spatial and temporal biases that must be taken into account. Instrumental records exist only for the last century or so, and thus proxy records can only be used to glean information about the climate for earlier time periods. But it must be noted that proxy records are not observations and strong caveats must be considered when they are used. It, too, must be noted that observational data are not without bias either.

Much research has described both the written and oral histories of the climate, as well as the proxy climate records. It is recognized that such records are not without their biases. For example, trees respond not to just air temperature fluctuations, but to the entire hydrologic cycle, including water supply, precipitation, and demand, which is only in part driven by air temperature.

Nevertheless, such accounts indicate that the climate of the last millennium has been characterized by considerable variability and that extended periods of cold and warmth existed. It has been gen-

erally agreed that during the early periods of the last millennium, air temperatures were warmer and that temperatures became cooler toward the middle of the millennium. This gave rise to the terms the Medieval Warm Period and the Little Ice Age, respectively. However, as these periods were not always consistently warm or cold, nor were the extremes geographically commensurate in time, such terms must be used with care.

In a change from its earlier reports, however, the Third Assessment Report of the Intergovernmental Panel on Climate Change, and now the U.S. National Assessment of Climate Change, both indicate that hemispheric and global air temperatures followed a curve developed by Dr. Mann and his colleagues in 1999. This curve exhibits two notable features, and I will point back to Dr. Mann's exhibit one that he showed a moment ago. First is a relatively flat and somewhat decreasing trend in air temperature that extends from 1000 A.D. to about 1900 A.D. This feature is an outlier that is in contravention to thousands of authors in the peer-reviewed literature.

This is followed by an abrupt rise in the air temperature during the 1900's that culminates in 1998 with the highest temperature on the graph. Virtually no uncertainty is assigned to the instrumental record of the last century. This conclusion reached by the IPCC and the National Assessment is that the 1990's was the warmest decade, with 1998 being the warmest year of the last millennium.

Despite the large uncertainty, the surprising lack of significant temperature variations in the record gives the impression that climate remained relatively unchanged throughout most of the last millennium, at least until human influences began to cause an abrupt increase in temperatures during the last century. Such characterization is a scientific outlier. Interestingly, Mann et al replace the proxy data for the 1900's by the instrumental record and present it with no uncertainty characterization. This, too, yields the false impression that the instrumental record is consistent with the proxy data and that it is error-free. It is neither.

The instrumental record contains numerous uncertainties, resulting from measurement errors, a lack of coverage over the world's oceans, and underrepresentation of mountainous and polar regions, as well as undeveloped nations and the presence of urbanization effects resulting from the growth of cities. As I stated before, the proxy records only in part reflect temperature. Therefore, a simultaneous presentation of the proxy and instrumental record is the scientific equivalent to calling apples and oranges the same fruit.

Even if a modest uncertainty of plus or minus one-tenth of a degree Celsius were imposed on the instrumental record, the claim of the 1990's being the warmest decade would immediately become questionable, as the uncertainty window would overlap with the uncertainty associated with earlier time periods. Note, too, that if the satellite temperature record, where little warming has been observed over the last 20 years, had been inserted instead of the instrumental record, it would be impossible to argue that the 1990's was the warmest decade. Such a cavalier treatment of scientific data can create scientific outliers, such as the Mann et al curve.

So we are left to question why the Mann et all curve seems to be at variance with the previous historical characterization of climatic variability. Investigating more than several hundred studies that have developed proxy records, we came to the conclusion that nearly all of these records show considerable fluctuations in air temperature over the last millennium. Please note that we did not reanalyze the proxy data. The original analysis from the various experts was left intact, as it formed a voluminous refereed scientific literature. Most records show the coldest period is commensurate with at least a portion of what is termed the Little Ice Age, and the warmest conditions at concomitant with at least a portion of what is termed the Medieval Warm Period.

Our conclusion is entirely consistent with conclusions reached by Drs. Bradley and Jones and not all locations on the globe experience cold or warm conditions. Moreover, we chose not to append the instrumental record, but to compare apples with apples and determine if the proxy records themselves indeed confirm the claim of the 1990's being the warmest decade of the last millennium. That claim is not borne out by the individual proxy records.

However, the IPCC report in the chapter with Dr. Mann as the lead author and his colleagues as contributing authors, also concludes that the research "support the idea that the 15th to 19th centuries were the coldest of the millennium over the northern hemisphere overall." Moreover, the IPCC report also concludes that the Mann and Jones research shows temperatures from the 11th to 14th centuries to be "warmer than those from the 15th to 19th centuries." This again is entirely consistent with our findings and in contravention of their own error assessment.

Where we differ with Dr. Mann and his colleagues is in the construction of the hemisphere average time series and their assertion that the 1990's was the warmest decade of the last millennium. Reasons why the Mann et al curve fails to retain the fidelity of the individual proxy records are detailed statistical issues into which I will not delve. But a real difference of opinion focuses solely on the Mann et al curve, and how it is an outlier compared to the balance of evidence on millennial climate change. In a very real sense, this is a fundamental issue that scientists must address before the Mann et al curve can be taken as fact.

In closing, let me state that climate is simply more than annually averaged global air temperature. Too much focus, I believe, has been placed on defining air temperature time series and such emphasis obscures the true issue in understanding climate change and variability. If we are truly to understand climate and its impacts and driving forces, we must push beyond the tendency to distill climate to a single annual number. Proxy records which provide our only possible link to the past are incomplete at best. But when these voluminous records are carefully and individually examined, one reaches the inescapable conclusion that climate variability has been a natural occurrence and especially so over the last millennium.

Given the uncertainties and biases associated with the proxy and instrumental records——

Senator INHOFE. Dr. Legates, we are going to have to cut it off. You have exceeded your time and I am sure you will have an op-

portunity to finish your thoughts during the question and answer period.

Dr. LEGATES. Thank you for the privilege.

Senator INHOFE. We are going to, if it is all right, use 5 minutes and maybe try to get a few rounds here. Is that acceptable? These will be 5 minute rounds for questioning. I will start.

First of all, Senator Thomas joined us. Thank you for coming, Senator Thomas.

I will address my first question to Dr. Legates. In my speech on the Senate floor yesterday, I noted your comments regarding—can you find that chart of those comments?—the comments regarding Dr. Mann's work as shown on the chart. I have a small copy of this. No, that is not it. It is this chart right here. OK.

First of all, this is a comparison. As I mentioned in my opening statement, we sit up here as non-scientists so we try to look at these things and see what is logical, how we should weigh and compare diverse opinions. Now, the first thing I noticed was that Dr. Mann, yours I believe was in the area of the timeframe of 1999——

Dr. MANN. Excuse me. That is incorrect.

Senator INHOFE [continuing]. And Dr. Soon, you are 2003. So I think that the timing would mean something because I know that this is not a static target. This is a moving target.

May I first ask Dr. Legates, do you stand by the statements that are made on this chart up here, on the contrasting methods that were used?

Dr. LEGATES. I have not had a chance to actually look at the chart before now.

Senator INHOFE. Is this the one that he had here? OK, let's put that up. All right, then, this statement here,

> "Although Mann's work is now widely used as proof of anthropogenic global warming. We have become concerned that such analysis is in direct contradiction to most of the research and written histories available. My paper shows this contradiction and argues that the results of Mann are out of step with the preponderance of the evidence."

I am not Tim Russert, but do you stand by these statements?

Dr. LEGATES. I do stand by them, sir.

Senator INHOFE. All right. I note that you are an expert in statistical techniques. In my speech on the Senate floor yesterday, I noted that even assuming all of the science used by the political left, come the end of 50 years hence, the Kyoto Protocol would have no measurable affect on temperature. Do you agree with that?

Dr. LEGATES. Yes, generally.

Senator INHOFE. And if the Kyoto Protocol forces harsher mandates, does it follow that the weaker legislative proposals that are out there right now before us in the Senate would have likewise no measurable effect?

Dr. LEGATES. That is likely true.

Senator INHOFE. All right. Let's see. Dr. Mann, since you have characterized your colleagues there in several different ways as nonsense, illegitimate, and inexperienced, let me ask you if you would use the same characterization of another person that I quoted on the floor yesterday. I would like to call your attention to the recent op/ed in the *Washington Post* by Dr. James Schles-

inger, who was Energy Secretary under President Carter. In it, he wrote, "There is an idea among the public that the science is settled. That remains far from the truth." He has also acknowledged the Medieval Warming Period and the Little Ice Age. Do you question the scientific integrity of Dr. Schlesinger?

Dr. MANN. I do not think I have questioned scientific integrity. I have questioned scientific expertise in the case of Drs. Willie Soon and David Legates with regard to issues of paleoclimate. As far as Schlesinger is concerned, I am not familiar with any peer-reviewed work that he has submitted to the scientific literature, so I would not be able to evaluate his comments in a similar way. If I could clarify one——

Senator INHOFE. OK. Well, you can't because there isn't time. I am going to stay within my timeframe and I want to get to questions so others will have plenty of opportunity to respond to questions I am sure.

Dr. Soon, how many studies did you examine in total and how many were appropriate for the criteria you established?

Dr. SOON. Senator, the number is roughly in the order of, if you speak in terms of the peer-reviewed literature, I would say several hundred. And the number of people involved in these paleoclimatic research would be at least 1,000. Of course, I have to emphasize I am not a paleoclimate scientist, but all of us are ruled by one simple goal, to understand the nature of how climate works. The basis to get to the goal is to figure out the exact expressions of the physical laws.

The short answer is there is a huge number of literature that we consulted that feed the criteria. This is why we wrote it as a scientific paper.

Senator INHOFE. I was trying to get to the 240 proxies that were used and the number used.

Dr. SOON. Yes, we listed about 240 proxy studies in our papers.

Senator INHOFE. Last, I would say, do you have more data in your study than Dr. Mann did in his 1999 work? And is your data newer?

Dr. SOON. Yes. I would emphasize that most of the proxy records come from the most recent 5 years.

Senator INHOFE. Thank you, Dr. Soon.

Senator Jeffords.

Senator JEFFORDS. Dr. Mann, would you care to respond?

Dr. MANN. Yes, first of all I wanted to clarify a misstatement earlier on the part of Senator Inhofe. The results that I showed in my first graphic which demonstrate that it is a clear consensus of the climate research community that a number of different estimates, not just ours, but at least 12 different estimates of the history of the northern hemisphere average temperature for the past 1,000 years give essentially the same result, within the uncertainties. We published a paper just a month ago demonstrating that that is a robust result of a large number of mainstream researchers in the climate research community.

Phil Jones and I also have a paper in press in the *Journal of Geophysical Research* letters, which demonstrates those results further. So in fact, the latest word and the word of the mainstream

climate research community is the one that I have given you earlier.

Now, as far as the issue of data, how much data was used, there are a number of misstatements that have been made about our study. One of them is with regard to how much data we used. We used literally hundreds of proxy records. We often represented those proxy records, as statistical climatologists often do, in what we call a state space. We represented them in terms of a smaller number of variables to capture the leading patterns of variability in the data. But we used hundreds of proxy indicators, more in fact than Dr. Soon referred to. In fact, we actually analyzed climate proxy records. Dr. Soon did not.

Senator JEFFORDS. Dr. Soon, in a 2001 article in Capitalism magazine, you said that because of the pattern of frequent and rapid changes in climate throughout the holocene period, we should not view the warming of the last 100 years as a unique event or as an indication of manmade emissions' effect on the climate.

But according to NOAA's Web site "upon close examination of these warm periods," including all the ones that you cited in your past and most recent article,

> "It became apparent that these periods are not similar to the 20th century warming for two specific reasons. One, the periods of hypothesized past warming do not appear to be global in extent or, two, the period of warmth can be explained by known natural climate forcing conditions that are uniquely different than those of the past 100 years."

Why didn't either of your articles make an impact on the state of the science or NOAA's position?

Dr. SOON. Thank you for your question, Senator. As you may be aware, my paper just got published this year, January 2003 and April 2003, so it is all fairly recent. I have just written up this paper very recently, so I do not know what impact it will have on any general community, but I do know all my works are done consulting works from all major paleoclimatologists in the field, including Dr. Mann and his esteemed colleagues.

As to the comments about the Capitalism magazine, I am not aware of that particular magazine. I do not know whether I submitted anything to this journal or this magazine. I do stand by the statement that it is important to look at the local and regional change before one takes global averages because climate tends to vary in very large swings in different parts of the world. That really is the essence of climate change and one ought to be really looking very carefully at the local and regional change first, and also one should not look strictly at only the temperature parameter, as Dr. Mann has claimed to have done. That I think is very important to take into account.

Senator JEFFORDS. Dr. Mann, could you comment?

Dr. MANN. Yes. Both of those statements are completely incorrect. If Dr. Soon had actually read any of the papers that we have published over the past 5 years or so, he would be aware of the fact that we use statistical techniques to reconstruct global patterns of surface temperature. We average those spatial patterns to estimate a northern hemisphere mean temperature, just as scientists today seek to estimate the northern hemisphere average temperature from a global network of thermometer measurements. We use pre-

cisely the same approach based on proxy reconstructions of spatial patterns of surface temperature.

So what Dr. Soon has said is completely inaccurate. The first line on that contrasting methods table up there is also completely inaccurate.

In terms of variables other than temperature, my colleagues and I have published several papers reconstructing continental drought over North America and reconstructed sea-level pressure patterns. We have looked at just about every variable that climatologists are interested in from the point of view of paleoclimate indicators. I think Dr. Soon needs to review my work more carefully.

Senator INHOFE. Thank you, Senator Jeffords.

Senator Allard.

Senator ALLARD. Thank you, Mr. Chairman.

In my mind, I do not think there is any question that the climate has shown a period of warming here. The question that I bring up and where I see the debate is, what is causing it and whether it is the changes that are happening and whether they are significant or not.

I also wonder what your thinking this world might look like 1,000 years from now, looking at the data that we have now. I wondered if maybe each one of you would just give me a brief response as to what you think of what we are seeing today may look like projected out over 1,000 years from now. I will start with Dr. Soon.

Dr. SOON. The factors causing climate change are extremely complicated. As I emphasized already, I am very much interested to learn how the climate changes on a local or regional scale first before I can speak in terms of global climate. After all, local and regional climate are indeed the most relevant climatic factors that human activities are being influenced by or the reverse way.

As to the factors of climate change, I believe that it is extremely difficult yet still to confirm the facts of being, let's say, even the late 20th century has anything to do with CO_2. We do know that the CO_2 is rising, but at the same time we know that climate depends on many other factors. It could be doing it internally all by itself because of ocean current movements. It could be done, for example, by variability imposed externally from the sun, variable outputs. Our sun is a variable star. That is a very well known fact.

These are the kinds of factors one has to look very comprehensively at. Additional important factors of human activity would include land use changes. Those are very well known factors that one has to keep a good record, or time history, to really understand what are the causes of the change.

I don't think I should speculate anything about futures. It is always very dangerous to talk about the future of any climate.

Senator ALLARD. Dr. Mann.

Dr. MANN. Yes. Well, I certainly agree with your statement that one of the key issues is what we call the detection or the attribution of human influence on climate, not just how has climate changed over the past 100 years or past 1,000 years, but can we actually determine the causal agents of change.

There has been a solid decade of research into precisely that question by, again, the mainstream climate research community in addressing the issue of the relative role of natural factors, as well

as anthropogenic factors. That includes the role of the sun, the role of human land use changes, and the role of human greenhouse gas increases. The model estimates are typically consistent with what we have seen in the observations earlier.

As far as the next 1,000 years, that is not a particular area of expertise of mine, but I am familiar with what the mainstream climate research community has to say about that. The latest model-based projections indicate a mean global temperature increase of anywhere between .6 and 2.2 degrees Centigrade. That is one degree to four degrees Fahrenheit relative to 1990 levels by the mid-21st century under most scenarios of future anthropogenic changes.

While these estimates are uncertain, even the lower value would take us well beyond any previous levels of warmth seen over at least the past couple of millennia. The magnitude of warmth, but perhaps more importantly the unprecedented rate of warming, is cause for concern.

Senator ALLARD. Dr. Legates.

Dr. LEGATES. Yes. I agree, too, that attribution is one of our important concerns. As a climatologist, I am very much interested in trying to figure out what drives climate. We know that a variety of factors exist. These include solar forcing functions; these include carbon dioxide in the atmosphere; these include biases associated with observational methods; these also include such things as land use changes. For example, if we change the albedo or reflected amount solar radiation, that too will change the surface temperature.

So it is really a difficult condition to try to balance all of these possible combinations and to try to take a very short instrumental record and discern to what extent that record is being driven by a variety of different combinations.

My conclusion probably in this case to directly answer your question is that the temperature likely would rise slightly, again due to carbon dioxide, but it would be much more responsive to solar output. If the sun should quiet down, for example, I would expect we would go into a cooling period.

Senator ALLARD. I guess the question that I would have, now, you know you have increased CO_2. So how is the environment in the Earth going to respond to increased CO_2? Have any of you talked to a botanist or anything to give you some idea of what happens when CO_2 increases in the atmosphere? Plants utilize CO_2, extract oxygen. We inhale oxygen and extract CO_2. Will plants be more prosperous with more CO_2? How does that impact the plant life? Can that then come back on the cycle and some century later mean more O_2 and less CO_2?

So I am wondering if any of you have reviewed some of these cycles with botanists and see if they have any scientific data on how plants respond to CO_2 when that is the sole factor. I am not sure I have ever seen a study. There is moisture and other things that affect plant growth, but just CO_2 by itself. Have any of you seen any scientific studies in that regard?

Dr. SOON. I have seen that. In fact, I have written a small paper that has a small section regarding that.

Senator ALLARD. And what was their conclusion?

Dr. SOON. The conclusion is that in general, of course, under enrichment of the CO_2 in the free air, that yes, plant growth will be enhanced. For example, as indicated by your chart, the crop yield can increase by 30 percent or higher for a doubling of CO_2, depending on the actual constraints in the field, like types of crops, how wet or how dry, etc. All of these examples are very well known and well verified in the field of botany.

Senator ALLARD. My time has run out. Would the other two agree with what he said?

Dr. MANN. Not quite.

Senator ALLARD. What is your modification?

Dr. MANN. In fact, a number of studies have been done, what are called "FACE" experiments. They are open canopy experiments in which CO_2 is elevated in the forest and scientists examine the changes in the behavior of that forest. What scientists at Duke University are finding is that while there is a tendency for an uptake of CO_2 by the plants in the near term, what happens is eventually those plants will die. They will rot. When that happens, this happens on generational time scales.

Senator ALLARD. Just CO_2 being the variable and not moisture and anything else?

Dr. MANN. Just CO_2. The CO_2 will go back into the atmosphere because the plants that take it up——

Senator ALLARD. Do they have an explanation of why the rot occurred?

Dr. MANN. Well, just when things die, they will rot and they will give up their CO_2 back to the atmosphere eventually.

Senator ALLARD. Well, that really does not get to the point I was trying to make.

Doctor.

Dr. LEGATES. To follow on that, enhanced CO_2 and dying plants would also provide the ability for more plants to therefore grow in its place. In particular, one of the people on our study, Dr. Sherwood Idso, has done a lot of this study with carbon dioxide and enhanced where you can control the amount of water and energy available to plants associated with lowered CO_2 and higher CO_2.

Senator ALLARD. So your conclusion is that CO_2 increases plant growth?

Dr. LEGATES. Yes.

Senator ALLARD. OK.

Thank you, Mr. Chairman.

Senator INHOFE. Thank you, Senator Allard.

Senator Carper, we were going to go by the early bird rule. Is it all right if Senator Thomas goes ahead of you here?

Senator CARPER. Sure.

Senator INHOFE. Senator Thomas.

Senator THOMAS. Thank you. I am a little confused about where we even ask the questions. Obviously, there is a difference of view. We are expected to make some policy decisions based on what we ought to be doing with regard to these kinds of things, but yet there does not seem to be a basis for that kind of a decision. Where would you suggest we get the information that is the best information we could get to make policy decisions for the future? Would each of you like to comment shortly on that?

Dr. MANN. Sure. I guess I would reiterate the comments that I made earlier, that in a National Academy of Sciences study that was commissioned by the Bush Administration in 2001, the National Academy of Sciences in essence stated their agreement with the major scientific findings of the Intergovernmental Panel on Climate Change, the IPCC, which is the United Nations panel of scientists, thousands of scientists from around the world who put together a report on the state of our knowledge about all of these things—climate change scenarios, our uncertainty about various attributes of the climate system. The conclusions that I stated earlier are the consensus conclusions of the IPCC.

Senator THOMAS. That is where you would go.

Dr. MANN. That is where they have gone, yes.

Dr. LEGATES. I would generally argue the IPCC is a bit of a political document to the extent to which it does present some biased science. There is a lot of good science in there, but a lot of the conclusions are sort of not borne out by the facts. Having been president of the Climate Specialty Group of the Association of American Geographers, which is probably the largest group of climatologists available, I know from talking to rank-and-file members that they generally—my impression is that most climatologists agree it takes a rather strong viewpoint.

So I have real serious concerns that it really represents a consensus, and in particular when, for example, in this discussion when we change dramatically what a lot of people have held true, that is the Little Ice Age, Medieval Warming and so forth, and replace it with a flat curve very quickly, I do not think we have given it enough time to really decide if in fact that is an appropriate change in paradigm.

Dr. SOON. Although I am not able to comment on anything on public policies, I am certainly able to testify that the science is completely unsettled. There are just so many things that we do not know about how the climate really works and what are the factors that cause it to change, to really jump to the conclusion that it will all be CO_2.

Senator THOMAS. Thank you. That helps a lot.

[Laughter.]

Senator INHOFE. You still have some time remaining. Did you have an opportunity to see the chart up here that Dr. John Reilly, MIT Joint Program on Science Policy and Global Change? On the floor yesterday, I talked at some length on this. There seems to be a lot of consensus that there are some very positive benefits.

Senator THOMAS. It is really interesting, you know, in Schlesinger's thing it indicates that the temperature after 1940 dropped until 1977. So that makes you wonder what we ought to do. The rise in temperature during the 20th century occurred between 1900 and 1940. So now we are faced with making policy decisions where there is no real evidence that the things that the greenhouse gases measurable by the U.N. is the basis for doing these things.

I know in science everyone has little different ideas, but I do think we are going to have to, Mr. Chairman, as you pointed out yesterday, either take it a little more slowly in terms of policy, or we are not going to have something more basic to base it on than we have now in order to make significant policy changes.

Thank you.

Senator INHOFE. Thank you, Senator Thomas.

Senator Carper.

Senator CARPER. Thank you, Mr. Chairman. I want to welcome our witnesses this morning. Dr. Legates, it is great to have a fighting Blue Hen here from the University of Delaware. We are delighted that you are here. Dr. Mann, thanks for coming up, and Dr. Soon, welcome. We thank you for your time and your interest and your expertise on these issues, and your willingness to help us on some tough public policy issues that we face.

Dr. Mann, I would start off if I could and direct a question to you. I understand we have had thermometers for less than 200 years, and yet we are trying to evaluate changes in temperature today in this century and the last century with those that occurred 500 or 1,000 or 2,000 years ago. I understand that we use proxies for thermometers, if you will, and for those kinds of changes in temperature.

I wonder if you could help me and maybe the committee better understand how we compare today's temperature measurements to the proxies of the past. Are there potential risks with relying on some of those proxies?

Dr. MANN. Absolutely. We have to use them carefully when we try to reconstruct the past temperature history. So when I say we have to use them carefully, it means some of the things that I discussed in my testimony earlier, that we need to actually verify that if we are using a proxy record to reconstruct past temperature patterns, that proxy record is indeed reflective of temperature changes. That is something that typically paleoclimate scientists first check to make sure that the data they are using are appropriate for the task at hand. Of course, we have done that in our work. I did not see evidence that Soon and colleagues have done that.

First of all, we next have to synthesize the information. There have been some misleading statements made here earlier on the part of the other testifiers with regard to local versus regional or global climate changes. Of course, we have to assimilate the information from the local scale to the larger scales, just as we do with any global estimate of quantity. So we take the regional information; we piece together what the regional patterns of change have been, which may amount to warming in certain areas and cooling in other areas. Only when we have reconstructed the true global or hemispheric regional patterns of change can we actually estimate the northern hemisphere average, for example.

A number of techniques have been developed in the climate research community for performing this kind of estimate. My colleagues and I have described various statistical approaches in the detailed climate literature. Some of the estimates are based on fairly sophisticated techniques. Some of them are based on fairly elementary techniques. Yet all of the results that have been published in the mainstream climate research community using different techniques and different assortments of proxy data have given, as I showed earlier in my graph, the same basic result within the uncertainties. That has not changed. An article that appeared last month in the American Geophysical Union, which is actually the

largest professional association of climatologists, showed that indeed that is the consensus viewpoint of the climate research community.

Senator CARPER. Thank you.

Dr. Legates, if I could ask a question of you, please. Have you or anyone of your colleagues, at the University of Delaware, to your knowledge studied the historical climate and temperature records in our part of the country, in Delaware, the Delmarva Peninsula, or the mid-Atlantic region?

Dr. LEGATES. We do not have anybody on staff presently that does paleoclimatology. One of the basic understandings that you must come up with when you study climate is that you must understand various things of hydroclimatology, physic climatology, and that includes paleoclimate study. So you must be at least versed in these things if you are not necessarily a paleoclimatologist.

We do have Dr. Brian Hanson at the University of Delaware who has looked at glacier movements over long time periods, as well as Dr. Fritz Nelson who has looked at changes associated with permafrost locations.

Senator CARPER. If someone were to do a study for our part of the country, what do you think they might find?

Dr. LEGATES. A study regarding?

Senator CARPER. Historical climate and temperature changes.

Dr. LEGATES. Over the East Coast of the United States? Most of the assessments indicated that generally the East Coast has gone through a variety of changes over long time periods. Historically, we have had a condition where in the 1960's, for example, we had conditions where there was much more snowfall. We have had a lot of variability associated with air temperature rising and falling over the local conditions. Variability is usually the characteristic of climate over the near-term as well.

Senator CARPER. OK. Dr. Soon, if I could ask you and maybe Dr. Legates the same question, the following question. That question is, do you believe that it is possible to emit unlimited amounts of CO_2 into our atmosphere without having any impact on climate or temperature?

Dr. SOON. I do not know how to precisely answer the question. If you fill up every single molecule of the air with CO_2, that would be poisonous, of course. I do not know the answer to the question, but I do like to add about the evidence available on climate change.

Senator CARPER. Before you do that, let me direct, if I could, the same question to Dr. Legates. I do appreciate your candor. It is not everyday that we find that here in this hall.

Dr. LEGATES. Generally, what we have found is that as carbon dioxide has increased, the temperature has followed, where in some cases historically the temperature has gone up and the carbon dioxide has fallen. So generally from a purely physical point of view, if you do increase the carbon dioxide, you should wind up with some trapping of gases, and hence wind up with a slightly increased temperature.

The question is, there is a lot of additional feedbacks associated with it. For example, warmer surface temperature leads to more instability or rising air which leads to more cloudiness. Clouds can

warm at night, but also reflect energy in the daylight. So you have these odd playbacks into the climate system which make it very difficult to say that if I hold everything else constant and change one variable, what will happen. Well, in reality, it is impossible to hold everything constant because it is a very intricate and interwoven system that one change does have feedbacks across the entire spectrum.

Senator CARPER. Thanks. I think my time has expired, Mr. Chairman. Is that correct?

Senator INHOFE. Yes. Thank you, Senator Carper.

Senator CARPER. Thank you.

Senator INHOFE. We will have another round here. In fact, I will start off with another round. Let's start with Dr. Legates. Dr. Legates, was the temperature warmer 4,000 to 7,000 years ago than it is today?

Dr. LEGATES. My understand was during about 4,000 to 7,000 years ago, in a period referred to as the climatic optimum, which sort of led to enhanced agriculture and led to development of civilization, generally the idea is that warmer temperatures lead to more enhanced human activity; colder temperatures tend to inhibit. Again, as we get back 4,000 to 7,000 years ago, it becomes, the error bars are getting wide as well. But the general consensus is that temperatures were a bit warmer during that time period.

Senator INHOFE. OK. Senator Thomas had something about, he had alluded to 1940. Yesterday when I was giving my talk and doing the research for that, it was my understanding that the amount of CO_2 emitted since the 1940's increased by about 80 percent. Yet that precipitated a period of time from about 1940 to 1975 of a cooling-off period. Is that correct?

Dr. LEGATES. That is correct. It is sort of a perplexing issue in the time series record that from 1940 to 1970 approximately, while carbon dioxide was in fact increasing, global temperatures appear to be decreasing.

Senator INHOFE. Dr. Mann, you have I might say impugned the integrity of your colleagues and a few other people during your presentation today. The Wharton Econometric Forecasting Associates did a study as to the effect of regulating CO_2 and what would happen. American consumers would face higher food, medical and housing costs; for food, an increase of 11 percent; medicine, an increase of 14 percent; and housing, an increase of 7 percent. At the same time, the average household of four would see its real income drop by $2,700 in 2010.

Under Kyoto, the energy and electricity prices would nearly double and gasoline prices would go up an additional 65 cents a gallon. I guess I would ask at this point, what is your opinion of the Wharton study?

Dr. MANN. OK. First, I would respectfully take issue with your statement that I have impugned the integrity of the other two testifiers here. I have questioned their, and I think rightfully, their qualifications to state the conclusions that they have stated. I provided some evidence of that.

Senator INHOFE. Well, "illegitimate, inexperienced, nonsense"——

Dr. MANN. Those are words that I used. Correct.

Senator INHOFE [continuing]. That is a matter of interpretation.

Go ahead.

Dr. MANN. I would furthermore point out that the very models that I have referred to track the actual instrumental warming and the slight cooling in the northern hemisphere. There was no cooling of the globe from 1940 to 1970, the northern hemisphere——

Senator INHOFE. OK. The question I am asking you is about WEFA.

Dr. MANN. I am not a specialist in public policy and I do not believe it would be useful for me to testify on that.

Senator INHOFE. Dr. Legates, have you looked at the report that Wharton came out with concerning the possible effects, economic results of this?

Dr. LEGATES. Again, I am not a public policy expert either, and so the economic impacts are not something which I would be qualified to testify on.

Senator INHOFE. OK, Dr. Legates, do you think you have more data than Dr. Mann?

Dr. LEGATES. I think we have looked at a large variety of time series. We have looked at essentially a large body of literature that existed both prior to Dr. Mann's analysis and since Dr. Mann's analysis, in attempting to figure out why his curve does not reflect the individual observations. It is one issue associated with when you put together data sets, to make sure that the composite sort of resembles the individual components.

Senator INHOFE. OK. The timeline, Dr. Mann, is something I have been concerned with, and those of us up here are listening to you and listening to all three of you and trying to analyze perhaps some of the data that you use and the conclusions you came to, having been 4 or 5 years back, compared to a study that was done referring to Smithsonian-Harvard, the 1,000-year study that was just completed, or at least given to us in March of this year. I would like to have each of you look at the chart up here and just give us a response as to what you feel in terms of the data that both sides are using today.

Dr. MANN. I guess you referred to me first?

Senator INHOFE. That is fine. Yes.

Dr. MANN. OK. Well, I think we have pretty much demonstrated that just about everything there is incorrect. In a peer-reviewed publication that was again published in the Journal *Eos* of the American Geophysical Union about a month ago, that article was cosigned by 12 of the leading United States and British climatologists and paleoclimatologists. We are already on record as pretty much pointing out that there is very little that is valid in any of the statements in that table. So I think I will just leave it at that.

Senator INHOFE. Do the other two of you agree with that?

Dr. LEGATES. If I may add, the *Eos* piece was actually not a refereed article. It is an *Eos* Forum piece, which by definition is an opinion piece by scientists for publication in *Eos*. That is what is contained on the AGU Web site for *Eos* Forum.

Senator INHOFE. All right. Let me ask one last question here. Dr. James Hansen of NASA, considered the father of global warming theory, said that the Kyoto Protocol "will have little affect on global temperatures in the 21st century." In a rather stunning followup, Hansen said it would take 30 Kyotos, let me repeat that, 30 Kyotos

to reduce warming to an acceptable level. If one Kyoto devastates the American economy, very much by the findings of Wharton, what would 30 Kyotos do? Is Dr. Hansen one of the most respected scientists in your field or is he way off base?

Dr. MANN. Dr. Hansen is certainly one of the most respected scientists in my field and I personally have great scientific respect for him. I think that his conclusions have been grossly taken out of context. His point is simply that Kyoto would, and this is his point, these are not my opinions, would do very little to ameliorate the warming over the next century for two reasons.

No. 1, there is something that scientists call the commitment to warming. Once we put CO_2 into the atmosphere, it takes many decades, on orders of decades to maybe centuries for it fully to equilibrate with the ocean and the atmosphere. So some of that CO_2 is taken up by the ocean. So the effect of it is delayed. So cutting back on CO_2 now may not affect global temperatures for 50 years, but 50 years later it is going to come back to roost.

Senator INHOFE. All right, that was a rather long answer, so let me just, with the indulgence of my fellow Senators here, I just want to ask one last question. I quoted Dr. Frederick Seitz, the past president of the National Academy of Sciences yesterday, and professor emeritus at Rockefeller University, who compiled an Oregon petition which says there is no convincing scientific evidence that human release of carbon dioxide, methane and other greenhouse gases is causing, or will in the foreseeable future cause catastrophic heating of the Earth's atmosphere and disruption of the Earth's climate.

Moreover, there is substantial scientific evidence that increases in atmospheric carbon dioxide produce many beneficial effects upon the natural plant and animal environments of the Earth. Do each of the three of you agree or disagree with his statement?

Dr. SOON. I agree.

Dr. MANN. I find little in there to agree with.

Dr. LEGATES. I would tend to agree.

Senator INHOFE. All right.

Senator Jeffords.

Senator JEFFORDS. As you may know, this is to all of you, the editor-in-chief of the magazine Climate Research resigned the position yesterday over problems with Dr. Soon's paper. In an e-mail sent to my staff, he said,

> "My view, which is shared by many, but not all editors and review editors of Climate Research, is that the review of the Soon et al paper failed to detect significant methodological flaws in the paper. The critique published in the *Eos* journal by Mann et al is valid. The paper should not have been published in this forum, not because of the eventual conclusion, but because of the insufficient evidence to draw this conclusion."

What methodological flaws does he mean?

Dr. Mann.

Dr. MANN. Well, I have tried to outline the most severe of those methodological flaws. I believe it is the mainstream view of just about every scientist in my field that I have talked to that there is little that is valid in that paper. They got just about everything wrong. They did not select the proxies properly. They did not actually analyze any data. They did not produce a reconstruction. They

did not produce uncertainties in a reconstruction. They did not compare to the proper baseline of the late-20th century in trying to make conclusions about modern warmth.

So I think it is the collective view of our entire research community that that is one of the most flawed papers that has appeared in the putative peer-reviewed research in recent years.

Senator JEFFORDS. Dr. Soon, do any scientists besides your co-authors support using wetness or dryness as indicators of past temperatures, instead of actual temperatures or proxy data that reflects temperatures?

Dr. SOON. As we explain clearly in our paper, and as it has been highly mischaracterized by my fellow colleague here, Dr. Mann, we certainly agree when we speak in term of the Medieval Warm Period, temperature is one of the important parameters. As we emphasize and specify in our papers that climate is not temperature alone. One has to look in terms of the water cycle, in terms of even the air cycles, in terms of the vegetation changes. These are the kind of details that we did not make any presumptions, but simply want to look at the patterns of change geographically all over the world, and see how complete the datas are, and then begin to start to see how do we assemble all such information.

Senator JEFFORDS. This is for the whole panel. I would like to know whether the unusual melting of Greenland ice sheets shown in this picture over the years 2001, 2002 and 2003, has been matched in the long-term climate history any other time? And according to NASA, by the end of the year 2002 season, the total area of surface melt in the Greenland ice sheet had broken all known records. By the end of that summer "Sea ice levels in the Arctic were the lowest in decades and possibly the lowest in several centuries."

NASA says this warming is happening faster and earlier than in previous periods. What is happening now and what is going to happen if this continues?

Dr. Mann.

Dr. MANN. Well, this is, of course, one particular region, one potentially isolated region, Greenland, in which there is evidence of mass oblation of ice. But if we look at what is going on the world over, mountain glaciers in the tropics throughout the world, glaciers in both the northern hemisphere and the southern hemisphere, what is seen is that glacial retreat during the late 20th century is unprecedented on similar time scales to the time scales I have spoken of before, the past 1,000 to 2,000 years.

I believe Professor Lonnie Thompson of Ohio State University has testified in this Senate before with regard to the dramatic evidence of worldwide glacier retreat. So that is a cause for concern. It is a harbinger of the warming because in fact the warming that is shown in those glacier retreats is actually warming that we are already committed to for decades to come.

Dr. LEGATES. Historically, it has been demonstrated in the refereed literature that much of this glacial retreat actually began in the late 1800's, before much of the carbon dioxide came into the atmosphere. This is very much consistent with the demise of the Little Ice Age and longer time-scale variations. Therefore, it is very difficult to say that these kind of events are directly attributable

to human impacts on the climate, when they in fact pre-date human impacts on the climate.

Senator JEFFORDS. Dr. Soon.

Dr. SOON. My only comment regarding that kind of chart or the claim that it has never happened before is that to think about the available, detailed observation that we have. We do not really have any satellite record longer than 20 to 30 years, so the statement that it has never happened before I think is dangerously inaccurate.

Senator JEFFORDS. Dr. Mann.

Dr. MANN. Yes. It is unfortunate to hear comments about the supposed inconsistencies of the satellite record voiced here, years after that has pretty much been debunked in the peer-reviewed literature, in Nature and Science. Both journals have in recent years published several rigorously peer-reviewed articles indicating that in fact the original statement that the satellite record showed cooling was flawed because the original author, John Christy, did not take into account a drift in the orbit of that satellite, which actually leads to a bias in the temperatures from the satellite.

Christy and colleagues have claimed to have gone back and fixed that problem, but just about every scientist who has looked at it says that their fix is not correct. If you fix it correctly, then the satellite record actually agrees with the surface record, indicating fairly dramatic rates of warming in the past two decades.

Senator JEFFORDS. I have one last question, Dr. Mann. What are the implications of your peer-reviewed work for future manmade warming?

Dr. MANN. As I said before, there have been a number of modeling simulations that have shown a fairly good match to our reconstruction and that of several independent research groups who have also produced these reconstructions of northern hemisphere temperature. So to the extent that the models match that record of the past 1,000 years when they are forced with various estimates of natural changes in the system, it gives us reason to trust what the models say about the future. As I testified before, the models tell us that we are likely to see a one degree to four degree Fahrenheit warming by the mid-20th century, given most predicted scenarios of continued anthropogenic influence on the climate.

Dr. LEGATES. If I may add something, one of the things I have heard is that science has been debunked and, for example, we pointed to Dr. Christy's curve up here and said that because one paper has been written, that curve is now called into question. We have talked about—you mentioned von Storch's resignation from Climate Research because apparently he has admitted that this paper never should have been published.

I want to point out that science debate goes on and on. In particular, Dr. Christy has had some very important contributions to indicate that his curve is not incorrect. That is part of scientific debate. Furthermore, I will say with respect to Climate Research, Otto Kinne, who is director of Inter-Research, the parent organization of Climate Research, asked Chris de Freitas who was the editor who served on the Soon and Baliunus papers, and I can relay this because I am a review editor of Climate Research so I am familiar with what has been taking place.

There were several people complaining that Chris de Freitas should be removed simply because he published the Soon and Baliunus paper. That question was brought to Otto Kinne. He asked for Chris de Freitas to provide him with the reviews, the changed manuscripts and so forth. He provided a letter in late June to all of us in which he said,

> "I have reviewed the evidence and I have indicated that the reviews, four for each manuscript, in fact there was a second or an earlier Soon and Baliunus article on another topic that was also called into question by these people leveling charges."

Essentially what he concluded was that the reviewers provided good and appropriate comments; that Doctors Soon and Baliunus provided an appropriate dressing or incorporation of these concerns; and that Chris de Freitas had in fact provided analysis appropriately.

Toward that end, Dr. von Storch was approached. Climate Research was putting in an editorial stating essentially this article should never have been published. Otto Kinne was informed and he has asked him not to submit that because it is not founded, and as a result Dr. von Storch, I now understand, has said he would resign.

Senator JEFFORDS. Dr. Mann.

Dr. MANN. Yes, just a very short comment. It is unprecedented in my career as a scientist to hear of a publisher of a journal going in and telling the editor-in-chief that he cannot publish an editorial. I find that shocking and a bit distressing. I do not know what the circumstances are behind it, but it is disturbing.

Dr. LEGATES. It is also unprecedented to find an editor being attacked, and this has also happened with the editorial staff of Energy and Environment, which is the other paper, to find an editor attacked for simply publishing an article that has been peer-reviewed and approved by reviewers.

Senator INHOFE. All right. The time has expired. We are 4 minutes over.

Senator JEFFORDS. I think that my witness should have the last word on my question, if I could. Dr. Mann, do you have any response to that?

Dr. MANN. Actually, my understanding is that Chris de Freitas, the individual in question, frequently publishes op/ed pieces in newspapers in New Zealand attacking IPCC and attacking Kyoto and attacking the work of mainstream climatologists in this area. So this is a fairly unusual editor that we are talking about.

Senator INHOFE. All right, thank you.

Senator Clinton has joined us. Senator Clinton would you like to have your round now?

Senator CLINTON. Thank you very much, Mr. Chairman. I thank you for this hearing. I understand that the questioning and the testimony has been somewhat lively, if not controversial and contested. The bottom line for me is whether we are doing what we need to do to ensure the best possible climatology outcome for future generations. I would stipulate that the Earth's climate has changed through the millennia. There is no doubt about that. I have read enough to know that we have had ice ages and we have had floods and we have had volcanoes. We have had lots of natu-

rally occurring events which have affected our climate. We have El Niño and his spouse, El Niña. We have all of that. That is not debatable.

The issue is whether the introduction and acceleration of anthropogenic activity primarily related to the burning of fossil fuels is putting into place conditions that will make it difficult, if not impossible for the Earth to regain its balance, that will support the conditions of life that we have inherited and are blessed with.

I know these debates have political implications because heaven forbid that we would tell somebody in the private sector not to do something, or that we might have to make sacrifices in the quality of our life for future generations. I think that it is not useful to carry out this kind of argumentation when it is clear that by the very nature of human development and industrialization, we have changed what is in the atmosphere, what is in the earth, what is in the waters.

That does not mean there was no change before we came along, and certainly in the last century that change has accelerated because the quality of life has improved, we have created chemicals that were never known in nature before. We have done a lot of things.

But I think that our goal should be to try to figure out how to do no harm or do the least amount of harm, and to ask ourselves, what are we willing to perhaps sacrifice to make sure that we are not contributing to irreversible changes. I know that academia is probably the most political environment in America. I was once on a staff of a law school. It was more difficult than any politics I had ever been involved in beforehand. I know that people have very strong opinions and hold on to them.

From my perspective, I just want to believe that I am making a contribution to ensuring that the quality of life for future generations is not demonstrably diminished. I would feel terrible if I participated, either as a willing actor or a bystander, in this potential undermining of our Earth's sustainability.

So Dr. Mann let me ask you, what was the Earth's climate like the last time that there was atmospheric concentration of carbon dioxide at today's levels of 370 parts per million?

Dr. MANN. Thank you, Senator, that is an excellent question. We have to go back fairly far into the past to find CO_2 levels approaching the CO_2 levels today. Ice core studies that have been done over the past decade or so have told us that today's CO_2 level is unprecedented now in at least four glacial or inter-glacial cycles. That is more than 400,000 years.

In fact, now as we look back from other evidence that is a bit more tentative, it appears that modern CO_2 levels probably have not been observed in 10 million to 20 million years. So we have to go back to the time of the dinosaurs, probably, to find CO_2 levels that we know were significantly higher than CO_2 levels today.

Some people will say, "Well look that was a great time." The dinosaurs were roaming near the poles. It was warm near the north pole. There were palm trees in the poles. Isn't that what we want? Well, that was a change that occurred on time-scales of tens of millions of years. What we are observing right now is a similar change that is occurring on time-scales of decades.

Senator CLINTON. Thank you. Thank you, Dr. Mann.

Senator INHOFE. Senator Clinton, if you would like to have some more time, since we are on the second round now, feel free to take another couple of minutes.

Senator CLINTON. Thank you very much, Mr. Chairman.

I guess that is, for me, the dilemma, because I certainly understand the testimony of the other two witnesses, and I read with great interest former Secretary Schlesinger's op/ed. I know that there are those, who are in a minority, let's at least admit that, who are in a minority, but who certainly have a very strongly held set of beliefs, and I respect that.

But I do believe that the compression of time in which these changes are occurring is extraordinarily significant. We can go back and look at the Earth's natural 125,000-year cycle, but I do not think we want to risk the enormous changes that could occur. I do not think we have a million or 10 million years or even 100,000 to experiment.

I think that the challenge confronting us is not to put our heads in the sand and let the academic argument take place, but figure out how in a sensible, prudent manner we could ameliorate these changes significantly enough so that if Dr. Soon and Dr. Legates are right, no harm done. If Dr. Mann is right, we will have saved ourselves a lot of potential damage and difficulty.

So I hope that we could put our heads together. I commend my two colleagues, both Senator Jeffords and Senator Carper, who have very sensible legislative answers to trying to get a handle on this. As I have said in this committee before, I stand ready to figure out ways to hold harmless our industrial base and others. I think it is a significant enough political, economic and moral challenge that if there are ways to make it financially possible for companies to do what needs to be done with respect to carbon dioxide and other atmospheric pollutants that have accelerated their presence in our atmosphere so dramatically in the last 100 years, I think we should do that.

This is not just a private sector problem. We all have benefited from the increasing use of fossil fuels, for example. Our standard of living is dramatically better. One of our problems is what is going to happen if China and India get a standard of living anywhere comparable to ours, and then begin to really—and I see Dr. Soon nodding—I mean really dump into the atmosphere untold amounts of new pollutants of whatever kind, leading certainly with carbon dioxide.

So this is a problem we need to get ahead of, and it is not a problem that the United States alone should be responsible for. It is not a problem that the private sector alone should be responsible for. But I believe, just as a prior generation of decisionmakers really put a lot of work into the law of the oceans and trying to figure out how we could protect our oceans, we need to do the same on the atmospheric level. There has got to be a way that we can come together on this big challenge.

So Mr. Chairman, I appreciate your continuing attention to this. I, for one, stand ready to work with you and our other colleagues because I just think this is too risky a proposition not to act on, given the weight of opinion, even with the dissenters, who I think

do rightly point out the incredible natural cycle, but we are now so influencing that natural cycle, I do not know if we have the time to contemplate the balance once again regaining itself in our wonderfully regenerating Earth.

Senator INHOFE. Thank you, Senator Clinton.

Senator Carper.

Senator CARPER. Thanks, Mr. Chairman. I just want to followup. Senator Clinton was kind in her comments on the legislation, the one that Senator Jeffords has introduced and second on legislation I have introduced along with Senators Judd Gregg, Lincoln Chafee and Lamar Alexander.

Are any of you familiar with that legislation? Would you like to become familiar over the next 5 minutes?

[Laughter.]

Dr. SOON. No, we will stick to science. Politics is too complicated.

Senator CARPER. All right. That may be the best approach.

We are trying to figure out if there is a reasonable middle ground on this issue. I am part of a group that Buddy MacKay, a former colleague of mine from Florida, calls the flaming moderates or flaming centrists. We can spend a whole lot of time discussing the impact of Kyoto caps, or we can focus on what steps we actually need to take.

The approach that Senators Gregg and Chafee and Alexander and myself have taken, at least with respect to four pollutants, we say unlike the President's proposal where he only addresses sulfur dioxide and nitrogen oxide and mercury, and does not address CO_2, as you know, because he thinks we need to study it a bit more. Our approach says that there ought to be caps on CO_2; that they should be phased in; that we should use a cap and trade system; we should give utilities the opportunity to buy credit for levels of CO_2 emissions that they maintain at high levels; and they should be able to contract with, among others, farmers and those who would be forced out of lands to change their planning patterns or change their animal feedlot operations in order to be able to sequester some of the CO_2 that occurs in our planet.

We have something called new source review. The President would eliminate it entirely. I think in Senator Jeffords' approach, it is pretty much left alone. There is a good argument that says that utilities under current law, if they make some kind of minor adjustment and minor investment in their plant, that they have to make a huge investment with respect to the environmental controls. As a result, it keeps them from making even common sense kinds of investments in their plants—sort of the laws of unintended consequences. That is sort of the approach that we have taken.

Now that you know all about it, if you were in our shoes, what kind of an approach would you take? Let me just start with our University of Delaware colleague here, Dr. Legates.

Dr. LEGATES. Generally, I favor no regrets policies, where they have other applications as well. But again, getting into the politics and the non-science aspects of what to do is out of my area of expertise. I may have my own beliefs, but they are no more important or less important than the average person. I would rather not testify to those here.

Senator CARPER. If you were convinced, and some of my colleagues have heard me talk about Dr. Thompson before, I don't know that they testified before this committee, but Doctors Knoll and Thompson spend their lives going around the world and they chart the disappearance of snow caps in some of the tallest mountains. I first met them here in Delaware about 5 or 6 years ago to receive an award for their research.

But they tell us that the snow caps around some of the tallest mountains in the world, the Himalayas and others, are not just disappearing, they will be gone, and they will be gone in our lifetime. When I heard them speak and talk about their work and what they were charting and finding, it got my attention. When you hear that, Dr. Legates and Dr. Soon, how does it affect you?

Dr. SOON. As a scientist, I am still questioning the actual evidence. The fact is that meltings may be recorded for certain glaciers. But among the things that we know is that there are about 160,000 glaciers around the Earth, but only 40 to 50 glaciers have been measured for 10 years or longer to tell us how much the ice has accumulated or has ablated.

Some of the specific melting examples, like Kilimanjaro, that Dr. Lonnie Thompson has looked at, or some places in Peru may be true. But the quality of the data records is really telling us that we do not have enough strong evidence to suggest that all the ice will disappear quickly and completely, or that all of it is unprecedented. Climate change is part of nature. As I tried to emphasize in my research by looking carefully into all the climate proxies, there are large local swings in the climatic changes.

Senator CARPER. Dr. Soon, what would it take to convince you that this is a problem we need to deal with?

Dr. SOON. As to some of the glaciers disappearing now in some parts of the mountains, I do not consider that to be either a problem or strong evidence——

Senator CARPER. No, no, the big issue. What would it take with respect to the concerns about global warming fed by CO_2 accumulation, what would it take to convince you that this is a problem we need to do something about?

Dr. SOON. OK. Scientifically, I would go by this very simple test. The simple test should be that the warming should be occurring first at the troposphere, the layer of air about four kilometers above us. That is a key part of the atmosphere that one should expect the CO_2 greenhouse effect to work its way downward toward the surface. I would urge, of course, very seriously that we do not lose sight in all these debates about science, we must sustain a certain kind of level of observational effort to keep track of data so that while we are arguing around what to do, that one has some records about any level of change that may occur.

So what it would take is that the CO_2 warming should happen at the layer of air four kilometers first. I would require it be strongly sustained for maybe 20 years or so. Then I would really believe that we have clear CO_2 fingerprints somewhere.

Senator CARPER. Mr. Chairman, I know my time has expired. Could I just ask that same question of Dr. Legates? What would it take to convince you?

Dr. LEGATES. Proof. Generally the problem we have seen in the record is that there is an awful lot of variability and there are things where changes occur, for example, between 1940 and 1970 where the temperature decreased, even though carbon dioxide was increasing. That sort of indicates to me that carbon dioxide may not be the biggest player in the game. Solar variability is likely to be the bigger player, changes in solar output. After all, if the sun goes out, our temperature drops considerably. We know historically that as the sun fluctuates in terms of its output, the climate does respond.

So there are a lot of other factors involved and I am not entirely convinced, based upon the proof, that carbon dioxide is a driving force. It is a contributory force in a small case, but not driving enough, because we wind up making policies potentially that can lead us to try to keep back the ocean, if you will. You cannot stop the waves from coming in.

Senator CARPER. Dr. Mann.

Dr. MANN. Two quick points. First of all, it grates on me to hear this argument about cooling from 1940 to 1970 continually cited here as evidence against anthropogenic climate change. That cooling was almost certainly anthropogenic and there has been a decade of research demonstrating that, anthropogenic sulphate aerosols, which have a cooling effect on the climate. What is happening now is that the much greater effect of increasing greenhouse gas concentrations is overtaking that small cooling effect of sulphate aerosols, also an anthropogenic influence, but not the one that is going to take us to doubled levels of CO_2 in the next century.

One quick other comment, if I could. Lonnie Thompson's work, which is some of the best work in our field, it is not like he has been looking for ice cores that are melting. He is actually looking for ice cores that are not melting because he wants to get long records. So if there is any belief that there might be some bias in the glaciers that he has gone to, if anything it is the opposite. He is looking for long records, so that makes it that much more impressive that they are all melting.

Senator CARPER. Thank you.

Senator INHOFE. Senator Allard.

Senator ALLARD. Thank you, Mr. Chairman.

What agency do you think we probably have the most expertise in as far as climatology change and what is happening with global climate? Would that be the agency on the National Oceanographic and Atmospheric Science, would that probably be where we would have most of our experts? If not, which agency do you think we would have most of our experts as far as the government is concerned? To any member of the panel, I would like to know whether any of you concur or not.

Dr. MANN. Well, I think that the different agencies specialize in different areas of the climate change research question, if you will. NOAA's specialty is in looking at climate variability, particularly with regard to oceanic variability. So they emphasize that area of the research. A lot of the peer-reviewed research, for example Lonnie Thompson's work that we just spoke of, is funded by the Na-

tional Science Foundation in large part. There are other organizations.

Senator ALLARD. The Foundation, is that an agency of the Federal Government?

Dr. MANN. Well, not directly.

Senator ALLARD. The question is, what is an agency of the Federal Government? The only one that I could think of was NOAA, but are there other agencies?

Dr. LEGATES. NASA does a lot of research, satellite-related efforts trying to estimate climate trends, incorporating satellite measurements as well.

Dr. MANN. As well as the Department of Energy and EPA.

Senator ALLARD. Yes, the Department of Energy.

Dr. LEGATES. The Department of Interior as well.

Senator ALLARD. OK. But we do not have any, say, each agency would have their own area of interest, but it seems to me that we need to look at global warming from a total perspective and I am trying to figure out if there is an agency that does that. I have talked to people within NOAA. There are arguments going on within that agency on the very topic that we are talking about here. There is absolutely no consensus within the agency, and I am trying to figure out if there is an agency out here that is taking on an overall view. I guess really there is not. We are just going to have to rely on the science community somehow or the other pulling all these views out from these various agencies. They look at the atmosphere, like you say, NASA looks at the stratosphere and higher up where your satellites are.

Dr. LEGATES. On the surface, too.

Senator ALLARD. We need somebody that looks at the effect on plant life, animal life, the total cycle; oxygen, CO_2 and all that before you reach conclusions. I am just wondering who pulls all this together so that we can come up with a total picture of what is happening as far as changes to this Earth is concerned, because it is more than just one science.

Dr. Mann.

Dr. MANN. There is a program, the U.S. Global Change Research Program, which seeks to coordinate the various agencies on issues of fundamental importance in the research of climate variability and climate change. So I think that is their role.

Senator ALLARD. OK. I want to get back a little bit to the absorption of sunlight, for example, on the Earth's surface. It seems to me, and I don't know how accurate this is. I want to check this out because it has been suggested to me by a number of people, that our absorptive surface on the Earth has increased. We still have the same amount of surface, but for example you have pavement in urban areas. We know that pavement is absorptive. Has that had an impact on global warming?

Dr. MANN. Most definitely.

Senator ALLARD. In your view?

Dr. MANN. Yes, your statement is correct. The main increase in the absorption by the Earth's surface is due to the melting of snow and ice. That has certainly had a very large influence on the warming, but it is part of the warming.

Senator ALLARD. So you do not think the construction of—we have more pavement than we did two centuries ago or a century ago.

Dr. MANN. Most models suggest that that is a cooling.

Senator ALLARD. Is there enough of that that we have more fields probably because of agriculture throughout the world, just not the United States. This is all over the world.

Dr. MANN. Yes. Most estimates suggest that there is a small cooling of the Earth's surface due to those changes.

Senator ALLARD. Would you all agree to that?

Dr. LEGATES. The pavements are associated with the urbanization effect, which is part of the problem associated with where we have observational measurements. Generally where you have a decrease in the light and heat exchange that is evaporation of water taking place because we have removed trees; the fact that you have darker surfaces; you have canyon-like effects. All of these lead to warmer temperatures in the city. The urban heat ion effect is well-documented and that is where virtually all of our observations are located.

But there are also changes in land surface effects by the fact that we are removing vegetation and replacing it with grasslands, for example, deforestation, de-vegetation. A lot of these are on very large-scales too, and they do change the color and character of the Earth's surface and hence the absorptive characteristic.

A lot of the cryosphere, a lot of the ice and snow is temporally variable. We have a growing area and decreasing area, so that does integrate itself out over time to some extent.

Senator ALLARD. Does the absorptive surface of the Earth's surface have an impact on whether we have a warmer temperature or not today?

Dr. LEGATES. Yes, absolutely.

Dr. SOON. Oh certainly, yes.

Senator ALLARD. I am a little bit confused of what the final view is. Do we increase temperature or do we cool the temperature?

Dr. MANN. Can I comment?

Senator ALLARD. Yes. You said that it cooled.

Dr. MANN. Yes, the effects that——

Senator ALLARD. OK, now, I would like to hear from——

Dr. Mann [continuing]. That is not the whole story. What he said is correct, but the effect that is dominant in models in about three or four different studies published in the past 2 years on precisely this question is actually the change in absorption by the land surface due to deforestation and other agricultural changes. That leads to an overall cooling of the globe, even in the face of other possible effects of warming.

Senator ALLARD. Would you agree with that?

Dr. LEGATES. Not necessarily. In particular, you are changing a characteristic, but you are also changing the other interactions. You are changing the vegetation and you are changing the evaporative characteristics.

Senator ALLARD. But your bottom line is that you think that, with increased absorptive rate on the Earth's surface, it has a cooling or a warming effect?

Dr. LEGATES. If you increase the absorption rate on the Earth's surface, you will have to have a net warming effect.

Dr. SOON. You have to have a warming.

Senator ALLARD. You have a warming.

I mean, to me this is a fairly fundamental concept, and here we are, we have disagreement at this table about that.

Dr. SOON. I don't think Dr. Mann is listening to your question.

Senator ALLARD. To me, from my practical experience, it seems to me that there is a warming effect. When I walk out on a pavement with my bare feet, they get burnt. If I walk on grass, my feet feel a lot cooler. I just look at it from a practical aspect. So Dr. Mann, would you explain to me why there is a difference in what you say and what I am feeling physically when I walk on the surface of the Earth?

Dr. MANN. Sure. When you are walking, you are only covering a pretty small fraction of the surface area of the Earth. The effect that you are talking about, for example, the urban heat island effect of blacktop and its tendency to absorb heat, that is overwhelmed by larger-scale changes that we do not necessarily see because they are not where we are walking around. Large areas of the surface area of the Earth are being changed in terms of their vegetation characteristics. That has a net cooling. The answer on that is clear in the peer-reviewed research.

Senator ALLARD. The reason I bring this up is that in the State of Colorado we have a lot of variation. We go from 3,000 to over 14,000 feet and we have a lot of different ecological systems in Colorado, depending on altitude and moisture and everything.

We have a weather reporting station in a rural area, in the plains of Colorado, and the data that I am getting from them, there is no indication of change as far as temperature is concerned. Yet as we move into the more urban areas, then we get weather stations that are indicating a higher temperature. So I am wondering worldwide, with the urbanization of the world, is there a possibility that we could be dealing with some temperature changes that are a result of the absorptive surface on the Earth like urbanization, you mentioned urbanization, we have a lot more than we used to have. Doesn't this have an impact on temperature?

Dr. LEGATES. Yes, definitely. Essentially, I do not think Dr. Mann answered the question appropriately in that your basic question was, if we absorb more radiation at the surface, will the temperature not go up? That is correct. The temperature will go up. In a sense, that is physics.

Senator ALLARD. Would you agree with that, Dr. Mann?

Dr. MANN. No. He has gotten about three different things wrong here.

Senator ALLARD. No, listen.

Dr. MANN. His first statement is wrong.

Senator ALLARD. I understand your statement. You are taking a broader atmospheric picture. You are taking a total picture. But the statement he made at this point, would you agree with that?

Dr. MANN. No. It is not correct.

Senator ALLARD. You would not agree?

Dr. MANN. The statement that he made was that there is an urban heat bias in the estimate of the surface temperature changes of the Earth.

Senator ALLARD. I did not hear him say that.

Dr. MANN. He said that earlier when he talked about urban heat bias.

Senator ALLARD. I am talking about the comment that he just made. Would you repeat the comment, Dr. Legates?

Dr. LEGATES. I essentially said the basic physics is that if you make the Earth's surface darker, you will absorb more energy, you will reflect less energy, as a result the surface temperature should increase.

Senator ALLARD. Would you agree with that scientific fact?

Dr. MANN. That statement would be in the first chapter of most textbooks. Yes.

Senator ALLARD. Dr. Soon, I did not mean to ignore you. You wanted to say something?

Dr. SOON. I tried to just emphasize that that is all you are asking.

Senator ALLARD. Yes.

Dr. SOON. If you increase absorptivity of the surfaces by changing it through any means, then more heat will be retained.

Senator ALLARD. I think part of the problem that we are running into here on the testimony is that we are not talking on the same terms. I think that we have to be very careful when we review the record and when we are listening to the witnesses here, Mr. Chairman, that we understand that we are all talking on the same terms in making the same point. I think the committee gets confused when we start talking from different terms and different perspectives.

I am just trying to simplify this argument down. I guess what I am coming to is that, as I have stated earlier, it is easy for me to believe that there is a trend in warming. The bottom line is what is causing it and what is going to be the long-term effects with this.

To me, the science is not entirely clear on that, and I do not see that that is being entirely clear on this panel because when I asked that question earlier, nobody gave me a specific on what they saw the effects were going to be. Maybe Dr. Mann did, and said that there was going to be warming. But most scientists when I talk to them just won't give me what they think the Earth is going to look like 1,000 years from now, or they will not necessarily step right out and say what are the causes of it because there are an awful lot of variables. I am not sure that scientists understand all those variables.

Dr. LEGATES. I think that is the issue. It is so uncertain and there are so many things that go into the mix, that to say fairly definitively it will be such in the future is very difficult to say.

Dr. SOON. We have to keep emphasizing that CO_2 is not the only player, the only factor. It is just highly short-sighted to just look at CO_2 as just one sole cause of change for every other change that we see or any variations that we manage to record.

Senator ALLARD. Yes. And when we talk about greenhouse gases, I think there is a tendency for us to think just in terms of CO_2.

Dr. SOON. Right.

Senator ALLARD. But isn't water vapor? Water vapor is a big part of greenhouse gases.

Dr. SOON. That would be the area of expertise by Professor David Legates. He studied that for almost 20 years.

Senator ALLARD. I do not know as we understand all of the aspects of each one of those fractionated, if we were to pull out each CO_2 or put out water vapor. What other gases do we have out there? Those are the main ones.

Dr. MANN. The other two have commented. May I comment as well?

Senator ALLARD. Let me finish my point. What are the greenhouse gases that we have?

Dr. MANN. I will speak to that.

Dr. SOON. Methane.

Senator ALLARD. Oh, methane. OK. We have methane. But the main ones are water vapor and CO_2. Water vapor being the largest, right?

Dr. SOON. Yes.

Dr. MANN. Can I comment on that?

Senator ALLARD. Dr. Mann.

Dr. MANN. Yes. There are trace gases like methane, carbon dioxide, chlorofluorocarbons, which we can actually control.

Senator ALLARD. Well, carbon dioxide is a very small part of greenhouse gases? Is that what you are saying?

Dr. MANN. No. There are several different greenhouse gases that we have to keep in mind, and it would be short-sighted to only talk about carbon dioxide. That is absolutely true.

Senator ALLARD. Right.

Dr. MANN. It is extremely misleading, however, when scientists cite the role of water vapor as a greenhouse gas. The concentration of water vapor in the atmosphere cannot be controlled by us directly, unlike the other trace gases. It is fixed by the surface temperature of the Earth itself. This is actually another chapter one textbook-type of result that we know to be true in the scientific community.

So we cannot change that freely. We can only change the other trace gases. When we do change those, we warm the Earth. We evaporate more water vapor and that gives us what we call a positive feedback that actually exaggerates the problem. But the water vapor itself cannot be the source of the problem.

Dr. SOON. It is really also scientifically inaccurate to say that we can really control CO_2. The global carbon cycle—we do not understand it well enough to really match or account for the CO_2 that we emitted. How much of it is really going into the ocean? How much of it has really gone into the forest? We do not have actually a full control of those parameters, as Dr. Mann would like to state on the record.

Senator ALLARD. Dr. Legates, do you have any comment?

Dr. LEGATES. Generally, the idea is that water vapor is the most important greenhouse gas. Period. That is Chapter One of any introductory text. The issue is, then, if we are associating with the effects of carbon dioxide and methane, which by the way has actually started to decrease over time, what we have found out is that

in particular we are dealing with with small matters where the bigger issues are not controllable.

Again, the sun is the biggest game in town and it is not controllable. At least I do not know that we can turn off the sun or control its output.

Senator ALLARD. OK. Senator Carper I think has a few questions.

Senator INHOFE. We have a serious problem here now, I am sorry to say, and that is that we are 30 minutes past our first panel and we are going to have to cut it off right now.

Senator ALLARD. OK, Mr. Chairman.

Senator INHOFE. I am very, very sorry. Thank you very much. I appreciate the fact that you are here.

We would call our next panel up. I apologize to the next panel because of the length of the first panel, we will have to cut this one short.

Dr. Leonard Levin is the program manager, Electric Power Research Institute; Dr. Gary Myers, professor of neurology and pediatrics, University of Rochester Medical Center; and Dr. Deborah Rice, the toxicologist, Maine Department of Environmental Protection, Bureau of Remediation and Waste Management.

I would like to ask each of you to confine your opening comments to 5 minutes, if you would. Your entire statement will be made a part of the record. We would start, Dr. Levin, with you.

STATEMENT OF LEONARD LEVIN, PROGRAM MANAGER, ELECTRIC POWER RESEARCH INSTITUTE

Dr. LEVIN. Thank you, Mr. Chairman, members of the committee.

I am Dr. Leonard Levin. I have come to discuss recent findings on mercury in the human environment. I serve as technical leader at EPRI, which is a nonprofit collaborative research organization. My remarks today represent my synthesis of research findings and are not an official statement of EPRI position.

It is a privilege to provide the committee this testimony on the science of mercury. I would like to address three key questions: sources of mercury; its deposition from the atmosphere to the Earth's surface; its potential accumulation in fish.

Where does mercury in the U.S. environment originate? Mercury is clearly a global issue. Recent estimates are that 2,340 tons of industry-related mercury are emitted globally. Over half of these originated from Asian sources. Of the global total, the United States is estimated to emit roughly 166 tons in total; U.S. utilities about 46 tons. In addition, it is estimated that another 1,300 tons of mercury emanates from land-based natural sources around the globe, and another 1,100 or so tons comes from the world's oceans.

Recent findings from the joint United States and Canadian METAALICUS field experiment show that a fairly small amount of deposited mercury, no more than 20 percent or so, re-admits to the atmosphere, even over a 2-year period. The implications are that mercury may be less mobile in the environment than we previously thought.

Studies by EPRI have shown that much of the mercury depositing in the United States may originate on other continents. Model

results show that for three-quarters of the continental U.S. land area, more than 60 percent of the mercury received comes from outside the country. Only 8 percent of U.S. territory receives two-thirds or more of its mercury from U.S. sources.

To check this with data, aircraft measurements were carried out by EPRI and the National Center for Atmospheric Research in Boulder, Colorado. Mercury and winds from the Shanghai, China region were tracked over the Pacific for 400 miles toward the United States. A second set of flights from Monterey, CA found that same plume from China crossing the California coast and entering U.S. territory. One implication is that there may be a management floor for U.S. mercury, a level below which the amount of mercury depositing to the surface cannot be reduced by domestic action alone.

Second, what are the primary sources of mercury in fish in the environment? Global mercury emissions appear to have peaked in the 1980's and declined or held steady since then. Professor Francois Morel of Princeton University, and colleagues, recently analyzed specific tuna for mercury, comparing recent catches with those from the 1970's. Despite changes in mercury emissions over those 30 years, mercury levels in tuna did not change between the samples. One conclusion they reached is that the mercury in such marine fish is not coming from emission sources on land, but from natural submarine sources of mercury. Again, this implies there may be a management floor for mercury in marine fish, which make up most of the U.S. fish diet.

Third, how can potential mercury reductions change mercury deposition? EPRI recently completed work to assess what might ensue in the atmosphere and in U.S. fish if further mercury emission reductions are carried out in the United States. The approach linked models of atmospheric mercury chemistry and physics with Federal data on mercury in fish in the U.S. diet, along with a model of costs that would be needed to attain a given reduction level. There are currently about 179 tons of mercury depositing each year in the United States from all sources, global and domestic. Current U.S. utility emissions of mercury are about 46 tons per year.

EPRI examined one proposed management scenario that cut these utility emissions from 46 tons to 25 tons per year. The analysis showed that this emissions cut of 47 percent resulted in an average 3 percent decline in mercury deposition in the United States. Some isolated locations making up less than one one-hundredth of the U.S. land area experienced drops of up to 30 percent. The economic model showed that costs to attain these lower levels would be between $2 billion and $5 billion per year for 12 years. This demonstrated U.S. mercury patterns may be relatively insensitive to the effects of this single category of sources.

In addition, most of the fish consumed in the United States are ocean fish which would be only slightly impacted by a reduction of 24 tons of mercury per year solely in the United States, out of 2,300 tons globally. Wild freshwater fish within the United States might show a greater reduction in mercury content, but they make up a very small part of the U.S. diet, compared to ocean or farm-raised fish.

These deposition changes were translated into how much less mercury might enter the U.S. diet via these three categories of fish. We found that less than one-tenth of 1 percent fewer children would be born at-risk due to their mother's taking in mercury at lower levels from fish consumed in the diet.

So to summarize, a drop of nearly half in utility mercury emissions resulted in an average drop of 3 percent in mercury depositing to the ground, and a drop of less than one-tenth of a percent in the number of children at risk. These recent findings are a small part of the massive international research effort to understand mercury and its impacts. EPRI and others, including U.S. EPA and the Department of Energy, are jointly racing to clarify the complex interactions of mercury with natural systems, an important part of its cycling, and its impacts on human health. With improved understanding, informed decisions can be made on the best ways to manage mercury.

Thank you for this opportunity to deliver these comments to the committee.

Senator INHOFE. Thank you, Dr. Levin.

Dr. Rice.

STATEMENT OF DEBORAH C. RICE, TOXICOLOGIST, BUREAU OF REMEDIATION AND WASTE MANAGEMENT, MAINE DEPARTMENT OF ENVIRONMENTAL PROTECTION

Dr. RICE. I would like to thank the committee for this opportunity to present information on the adverse health consequences of exposure to methyl-mercury in the United States.

I am a neurotoxicologist who has worked on the neurotoxicity of methyl-mercury for over two decades and have published over 100 papers on the neurotoxicity of environmental chemicals. Until 3 months ago, I was a senior toxicologist at the Environmental Protection Agency. I am a coauthor of the document that reviewed the scientific evidence on the health effects of methyl-mercury for EPA. This document included the derivation of the acceptable daily intake level for methyl-mercury.

I would like to focus on four points. No. 1, there is unequivocal evidence that methyl-mercury harms the developing human brain. No. 2, EPA used analyses of three large studies in its derivation of an acceptable daily intake, including the studies in the Seychelles Islands which found no adverse effects. No. 3, 8 percent of women of childbearing age have levels of methyl-mercury in their bodies above this acceptable level, and studies have documented cardiovascular disease in men at low levels of methyl-mercury, suggesting that an additional potentially large segment of the population is at risk.

Studies performed around the world have documented harmful effects of environmental methyl-mercury exposure on children's mental development. Three major studies were analyzed by the National Research Council panel in their expert review: In the Faroe Islands in the North Atlantic, and the Seychelles Islands in the Indian Ocean, and in New Zealand. Two of these major studies, as well as six smaller studies, identified impairment associated with methyl-mercury exposure. The Seychelles Island study is anomalous in finding no effects. Adverse effects include decreased IQ and

deficits in memory, language processing, attention and fine motor coordination.

The NRC modeled the relationship between the amount of methyl-mercury in the mother's body and the performance of the child, and calculated the level associated with the doubling of the number of children that would perform in the abnormally low range. The NRC panel did this for each study separately and for all of the three studies combined, including the negative Seychelles study.

EPA used the NRC analyses in deriving its acceptable daily intake level of methyl-mercury. EPA performed the relevant calculations based on each of the two positive studies, as well as the integrative analysis of all three studies. The acceptable level is the same whether it is based on the integrative analysis of all three studies, or on the Faroe Islands study alone.

The acceptable level would be lower if only the New Zealand study were considered. Only if the negative Seychelles study alone were used, while ignoring the values calculated for the Faroe Islands and New Zealand studies, would the acceptable intake level be higher than the current value. EPA believed that to do so would be scientifically unsound and would provide insufficient protection to Americans.

Data from a survey representing the U.S. population collected over the last 2 years revealed that about 8 percent of women of childbearing age had blood concentration of methyl-mercury above the level that EPA believes is safe. This translates into over 300,000 newborns at risk for adverse effects on intelligence and memory, ability to pay attention, language skills and other abilities that are required to be successful in our highly technological society.

There is an additional concern regarding the potential for harm as a result of environmental methyl-mercury exposure. Three studies found a relationship between increased methyl-mercury levels and atherosclerosis, heart attacks and death, and it is unknown whether there is a level of mercury that will not produce harm. It is important to understand that the cardiovascular effects associated with methyl-mercury may put an additional very large portion of the population at risk.

In summary, there are four points that I would like the committee to keep in mind. First, at least eight studies based on populations around the globe found an association between methyl-mercury levels and impaired neuropsychological function in children. The Seychelles Islands study is anomalous in finding no effects. Second, both the NRC and the EPA included the Seychelles Islands study in their analysis. The only way that the acceptable intake of methyl-mercury could be higher would be to ignore the two major positive studies, as well as six smaller studies and rely solely on the one study that showed no effects.

Third, there is a substantial percentage of women of reproductive age in the United States with levels of methyl-mercury in their bodies above what EPA considers safe. As a result, over 300,000 newborns each year are exposed to potentially harmful levels of methyl-mercury. Fourth, increased exposure to methyl-mercury may result in cardiovascular disease and even death in men from

heart attack, suggesting an additional large segment of the population is at risk.

Additional information has been provided to the committee. Thank you for your time and attention.

Senator INHOFE. Thank you, Dr. Rice.

Dr. Myers.

STATEMENT OF GARY MYERS, PROFESSOR OF NEUROLOGY AND PEDIATRICS, DEPARTMENT OF NEUROLOGY, UNIVERSITY OF ROCHESTER MEDICAL CENTER

Dr. MYERS. Thank you for the opportunity to present the views of our research group on the health effects of methyl-mercury exposure. My name is Gary Myers. I am a pediatric neurologist and a professor at the University of Rochester in New York, and just one member of a large international team that has been studying the human health effects of methyl-mercury for nearly 30 years. For 20 of those years, our group has specifically studied the effects of prenatal methyl-mercury exposure.

In 1971 and 1972, there was an epidemic of methyl-mercury poisoning in Iraq. The source of exposure, unlike in Japan, was maternal consumption of sea grain coated with a methyl-mercury fungicide. We looked at a number of children in that study and measured the exposure of the fetus using the maternal hair as the biomarker. It is the only biomarker that has been correlated with brain levels. We concluded that there was a possibility that exposure as low as 10 parts per million in maternal hair might be associated with adverse effects on the fetus. This value is over 10 times the average in the United States and five times the average in Japan, but individuals consuming large quantities of fish can easily achieve this level.

The hypothesis of our study in the Seychelles was that methyl-mercury from fish consumption might affect child development. In fact, we all thought it would. Since millions of people around the world consume fish as their primary source of protein, we thought it was only reasonable to investigate the question directly. We selected the Seychelles because of two reasons. First, they eat large amounts of fish. The average mother eats 10 times as much as women here in the United States.

Second, the fish in Seychelles has an average mercury content of about 0.3 parts per million, which is approximately the same as commercial fish here in the United States. The Seychelles study is a collaborative study which was begun under the auspices of the WHO and has been carried out by a U.S.-led team of international researchers from the University of Rochester, Cornell University and the Ministries of Health and Education in Seychelles. The funding has come from the National Institutes of Environmental Health Sciences, with some minor funding from the Food and Drug Administration and the governments of Seychelles and Sweden.

The Seychelles was chosen for a number of reasons, primarily because there was no overt mercury pollution and many of the factors that complicate epidemiological studies of low-level exposures were simply not present. There was universal free and readily available health care in Seychelles. Prenatal care is nearly 100 percent. The birthrate is high and the general health of the mothers and chil-

dren is very good. In addition, education is free, universal, and it starts at age 3½.

Before starting the study, we carefully controlled for a number of things. To minimize the possibility of bias, a number of decisions were made. First, no one in Seychelles, including any of the researchers who visit the island, would know the level of exposure of any child or mother unless our results indicated that children were indeed at risk. Second, because of the known problems with developmental delay in certain disorders, those children would be excluded from the study. Third, the tests administered would include all of the tests that have been used in other studies, plus other things that we thought might detect subtle changes.

Fourth, we would do this testing at specific age windows. Fifth, we would adjust for multiple confounding factors, things that are actually known to affect child development such as socioeconomic status, the mother's intelligence, and birth weight. And sixth, we established a data analysis plan before the data were collected to minimize the possibility that the data would just be repeatedly analyzed until the anticipated effect was in fact determined.

We have now carried out five evaluations of the children over 9 years. The study has focused on prenatal exposure. The exposure of both mothers and children has been in the range of concern, from 1 to 27 parts per million. We have done extensive testing with over 57 primary endpoints determined so far. The study has found three statistical associations with prenatal methyl-mercury exposure. One was adverse; one was beneficial; and one was indeterminate. These results might be expected to occur by chance and do not support the hypothesis that adverse developmental effects result from prenatal methyl-mercury exposure in the range commonly achieved by consuming large amounts of fish.

The findings from our research have been published in the world's leading medical journals, including the Journal of the American Medical Association, the Lancet, and a soon-to-be-published review in the New England Journal of Medicine. We do not believe that there is presently good scientific evidence that moderate fish consumption is harmful to the fetus. In the words of Dr. Lyketsos, a distinguished researcher from Johns Hopkins, who wrote the editorial with our Lancet articles:

> "On balance, the evidence suggests that methyl-mercury exposure from fish consumption during pregnancy of the levels seen in most parts of the world does not have measurable cognitive or behavioral effects in later childhood. However, fish is an important source of protein in many countries and large numbers of mothers around the world rely on fish for proper nutrition. Good maternal nutrition is essential to the baby's health."

Thank you.

Senator INHOFE. Thank you, Dr. Myers.

We are going to try to adhere to a 5-minute round of questioning. Let me just share with you, which I think you already know, you folks are looking at the medical effects of mercury. We also up here have to consider the economic effects—the problems that are out there. Right now on the Senate floor, they are debating the energy bill. We have an energy crisis in this country, and if cofire should go out, and that could happen from either CO_2 or mercury, it would be a very serious crisis. I think anticipating that this will happen,

several people have moved off-shore, moved to other places. So that is something that is really, I guess you would say our major, at least one of my major concerns.

Now, just for all of the witnesses, you stated that the U.S. utility mercury emissions are 46 tons a year. Tell us what happens to this mercury. Help us visualize where does it come from; where does it go; how much is deposited in the United States; how does this compare with the amount that is deposited in the United States from global sources.

Would you like to start, Dr. Rice?

Dr. RICE. That is really not my area of expertise, so I cannot speak to it.

Senator INHOFE. All right.

Dr. Myers.

Dr. MYERS. It is not my area of expertise.

Senator INHOFE. Come on, Dr. Levin.

Dr. LEVIN. All right.

[Laughter.]

Dr. LEVIN. Utility mercury of the various sources of mercury is probably the best-studied category, partially because there are more individual sources than there are of many of the other categories. We believe that roughly half on average coming out from utilities is made up of the divalent form of mercury, which is about a million times or so more soluble in water than the elemental form, which is the silvery liquid that you probably remember from high school chemistry. So of this mercury emitting from all utilities in the United States, roughly half of it is more highly water soluble and the other half will tend to go into regional and global circulation.

We calculate that about 70 percent or so of the mercury emitted from utilities leaves the United States, and the other 30 percent or so deposits within the United States across the country. These are somewhat similar to the numbers that EPA is deriving as well. Some of this mercury that deposits to the surface will wind up in receiving waters, and a very small fraction of it, probably less than 1 percent, will eventually be turned into the organic form by bacterial action. It is that organic form that has the potential to reach humans through accumulation in some fish.

Again this does not happen in all waterways and with all fish species. It tends to happen in waterways that have full food webs that go to high-level fish that grow quite large, and it is larger, older fish that tend to accumulate more mercury.

Of the exposure in the community in the United States, almost all of it is through intake from fish and the mercury in those fish, although the levels taken in can vary from very little or almost none, to amounts of concern. There is almost no exposure by inhalation. That is a very small part of the exposure.

So our concern is to follow this mercury from its sources through to where it winds up in fish and eventually may be consumed by humans. That is the trick, scientifically.

Senator INHOFE. Thank you, Dr. Levin.

Dr. Rice, the American Heart Association and the World Health Organization recommend that fish should be a part of everyone's diet, concluding that the benefits of eating fish outweigh the risks

of adverse effects, which as you state in your testimony are potential risks. Since eating fish offers substantial health benefits, shouldn't the EPA's referenced dose be revised to take this into account, or does it?

Dr. RICE. Well, I agree totally, and I have to say that I am no longer with EPA so I am not speaking as a representative of the agency. I need to make that clear. So some of these opinions will be those of the agency when I left, and some will be mine.

But the scientific community at large and the EPA and me personally recognize that fish is a good source of protein. It also confers cardio-protective effects. There are also omega-three fatty acids in fish that are essential when the fetus is building its brain. There is new evidence that eating fish also may be beneficial to the mental development or the mental function of the elderly. I suspect that it is probably important for all of us.

So the dichotomy is not eat fish/don't eat fish. The important thing to be able to do is to come out with some recommendations to the community that allow people to eat fish, but not to eat fish that has increased levels of methyl-mercury. So EPA thinks that, I was part of that EPA panel, so when I was part of that EPA panel we firmly believe that the RFD should not be any higher, and in the light of some evidence that we were not able to analyze at the time, might even should be lower than it is presently.

So it is not a question of increasing the reference dose. It is a question of making sure that the American public can eat fish that does not have undue levels of methyl-mercury in them.

Senator INHOFE. Thank you very much.

Dr. Myers, in selecting the Seychelles as a location for your research, what other locations did you consider other than the Seychelles Islands?

Dr. MYERS. We started studies on the coast of South America and looked also at the Maldive Islands as another possibility.

Senator INHOFE. Yes. I kind of wanted to get to the Faroe Islands. Did you consider them for your research?

Dr. MYERS. We did not consider the Faroes in our research.

Senator INHOFE. It is my understanding that, and for those of us who are not scientists here, that some of the problems, let's take the Faroe Islands and see if I have this right, that there is an inordinate amount of whale meat that is consumed there and there are PCBs in there. I do not know whether you can distinguish between the harm of one or the other, but is this a factor that should be considered?

It is my understanding, and I won't say this right, but there are different levels of mercury that are found. One is from the primary fish, and the other is from whales that eat other fish, so it has a multiplying effect. Is this taken into consideration?

Dr. RICE. The Faroe Islands study and the Seychelles Islands together have been reviewed by at least two very distinguished peer-review panels. That issue, the issue of the pattern of intake of methyl-mercury and potential co-exposure for PCBs has been discussed extensively by the scientific community.

The Faroe Islands' population does eat whale meat. They may eat a large whale dinner occasionally. They also tend to dry the whale meat, and so they snack on it in addition to eating a so-

called bolus dose, what we call a bolus dose. So they have a low level of methyl-mercury intake which may be occasionally punctuated with a higher intake level. The source of methyl-mercury does not matter, whether it is through fish or through whale. So the fact that it is whale meat per se is not really relevant.

None of the panels, including the National Research Council panel, could come to any kind of conclusion about the importance of the pattern of intake, because the data just are not available. There just are not scientific data that speak directly to that. But what the Faroe Islands investigators have done because this was raised as a concern and because they have hair, and they had hair from their population that was stored, they were able to go back and do segmental analysis, so that you cut the hair up into tiny little pieces and look at mercury levels across the length of the hair.

What they did was they eliminated the mothers that had the most variable hair levels that might suggest that there was this bolus exposure of these particular women and these particular fetuses. What they found was that the effect was actually stronger when they eliminated these women, which makes a certain amount of sense because you are decreasing variability when you do that.

Senator INHOFE. Thank you, Dr. Rice.

Senator Jeffords.

Senator JEFFORDS. Thank you all for your testimony on this very important and timely topic.

Some of you have seen this morning's New York Times full-page article on mercury and its health effects. This helps to set a context for our discussion.

Dr. Rice, what exactly is a reference dose level and what does it mean in terms of the so-called safe levels of fish consumption? Does EPA reference dose level include a built-in tenfold safety threshold?

Dr. RICE. The reference dose is designed to be a daily intake level that a person could consume over the course of their lifetime without deleterious effects. So it is designed to be the amount of mercury you could eat every day in your life and not harm yourself.

Now, when EPA did its calculation, it is important to understand that when the National Academy of Sciences modeled a number of endpoints for each of the studies, and those were the Faroe Islands study, the New Zealand Study, both of which found effects, as well as the Seychelles study which did not, they identified not a no-effect level. They identified a very specific effect level. That effect level is associated with a doubling of the number of children that would perform in the abnormal range, in other words, the lowest 5 percent of the population. So this is in no way a no-effect level.

To that, the EPA applied a tenfold so-called uncertainty factor. The point of that was to take into account things that we did not know, data that we did not have, as well as the pharmacodynamic and the pharmacokinetic variability. Now, there were actually data that was again modeled by the NAS and reviewed by the NAS, that says that the pharmacokinetic variability, in other words the woman's ability to get rid of methyl-mercury from her body, differs by a factor of three. So that already takes up half of the uncertainty factor.

But in addition to that, it is important to understand that when the Faroe Islands folks analyzed their data, they eliminated mothers with mercury levels above 10 ppm in their hair, which was really right about at the effect level that the NAS identified. The effects were just about as strong even below 10 ppms. So again, that is very strong evidence that there is not a factor of 10 safety.

In addition to that, when the NAS modeled their data, it turned out that both of the New Zealand study and the Faroe Islands study not only was there no evidence that there was a threshold, in other words a level below which there were no effects, but in fact the curve was actually steeper at the lower levels. The NAS used a straight line when they modeled the data because they were uncomfortable about using curves that were steeper at the lower end than they were at the higher end, but subsequent to that there have been studies come out with regard to lead exposure, for example. There are now several studies where that has also been found for lead exposure.

So this may in fact be a very real effect. So not only is there not a safety factor of 10. There might be virtually no safety factor at all.

In addition to that, something that EPA recognized at the time, but we were not able to quantitate because we did not have the data, but it has now been quantitated, we assumed that the relationship between the mother's blood level of methyl-mercury and the fetus' blood level of methyl-mercury were the same, because of course we have the body burden; we have cord blood in the fetus, we have to get back to intake by the mother. We know now that in fact the ratio is more like 1.7, and for some mothers it is as much as over 3.

So if we were to recalculate the reference dose just based on this new information, it would decrease from 0.1 to 0.06.

Senator JEFFORDS. Dr. Rice and Dr. Myers, would you recommend that Members of Congress and regulatory agencies base their decisions on whether and how much to reduce human-made mercury emissions on the findings from any one study?

Dr. MYERS. Our group has been involved in the science of studying whether you could find effects at low levels, and we have not been involved in policy. There is a general scientific principle, I think it is important to look at multiple different studies. However, these studies are complicated and one has to look at what kind of studies you are dealing with. Some are simply descriptive. They take a group of people and describe something. It is a basic epidemiological principle that you cannot assign causation from a descriptive study.

So one has to look at the studies that are larger and follow children over time, and control for a lot of confounding factors which complicate these type of studies very much actually. The Seychelles study in fact is not a negative study, as has been stated. We did, in fact, find associations with things that are known to affect child development, such as socioeconomic status, maternal intelligence, the home environment and other things. What we did not find was an adverse association with prenatal methyl-mercury exposure in the Seychelles.

Senator JEFFORDS. Dr. Rice.

Dr. RICE. I agree with Dr. Myers. These studies are very complex. I think that that is even more reason not to rely on one study while eliminating other studies for consideration.

Again, these studies have been peer-reviewed numerous times. The Seychelles Islands study and the Faroe Islands study have been reviewed now by several panels. They are both thought to be very high quality, very well-designed and well-executed studies.

The NAS, as well as the previous panel, talked at great length about what might account for the differences between these studies. We really do not know what accounts for the differences between these studies. The NAS modeled three studies. The New Zealand study was also a positive study.

The National Academy of Sciences and the EPA agreed with them that it was not scientifically justifiable for protection of the health of the American public to rely on the negative study and exclude the two positive studies. I said at least a couple of times in my testimony that what the NAS did to try to address that was to do an integrative analysis that included all three studies, including the Seychelles Islands study, and modeled it statistically.

When EPA then took those analyses and derived, what we did was we derived a series of reference doses, kind of sample reference doses, that were based on a number of endpoints from both the New Zealand study and the Faroe study, as well as the integrative analysis of all three studies. The integrative analysis of all three studies also yields a reference dose of 0.1. So that made me personally very comfortable that we were doing the right thing scientifically in our derivation of the reference dose.

Senator INHOFE. These are supposed to be 5-minute rounds and it has been 8 minutes, so we will recognize Senator Allard.

Senator ALLARD. Dr. Rice and Dr. Myers, you have in your comments talked about methyl-mercury as being the toxic compound as far as human health is concerned. Are there other mercurial compounds that are toxic to humans?

Dr. RICE. Yes. All forms of mercury are toxic to humans.

Senator ALLARD. Including the elemental form?

Dr. RICE. Yes.

Senator ALLARD. OK.

Dr. RICE. But in terms of environmental exposure, it is really the methyl-mercury form that we are worried about because that is the form that gets into the food chain and is concentrated and accumulated up the food chain. That is what people actually end up being exposed to.

Senator ALLARD. OK. Thanks for clarifying that. I appreciate that. So this gets into the environment and consequently in the fish or food chain or whatever. Is the starting point always bacteria operating on the elemental form of mercury? Or is it these various compounds that bacteria operate on and then end up being assimilated into the food chain? How does that happen?

Dr. RICE. In most circumstances, it is the inorganic form, not the elemental mercury, but the inorganic form that is available to be taken up by various microorganisms.

Senator ALLARD. How do we get to that organic form, the methyl-mercury? How do we get there?

Dr. RICE. The microorganisms actually put a methyl group on as part of their metabolic processes.

Senator ALLARD. Do they get that from elementary mercury? Is that the origin, or is it various compounds of mercury?

Dr. RICE. Yes, it is just straight mercury. Now, in the Japanese outbreak, it was actually methyl-mercury that was put into the water, but that is a relatively unusual situation.

Senator ALLARD. I see. OK, so my understanding, Dr. Levin, is that a lot of the mercury that is introduced into the environment of this country does not originate within the borders of this country. Is that correct? The suggestion is that a lot of the sources of mercury that come across that we may pick up in the soil is actually carried over by wind and what not from the Asian countries. Is that correct?

Dr. LEVIN. That is correct, Senator, as far as the modeling shows, and that is consistent with work that EPRI has done, EPA and others have also done in the modeling.

Senator ALLARD. Is this the elemental mercury that is being brought over?

Dr. LEVIN. It is elemental, or the elemental form. It is also the inorganic form or the form that can be combined into salts.

Senator ALLARD. Now, the inorganic form is not processed into the food chain? Did I understand that correctly?

Dr. LEVIN. It is the inorganic form that is processed into the food chain.

Senator ALLARD. Yes, it is the organic form.

Dr. LEVIN. The two forms that are emitted from combustion sources are the elemental form, the chemicals found on the periodic chart.

Senator ALLARD. Right.

Dr. LEVIN. And the inorganic form, which combines with, for example, chlorine, to form the pure chloride, or is the form also found in minerals. Those two forms that wind up in the proper aquatic environments, it is the inorganic form that may be methylated and turned into the organic form.

Senator ALLARD. Right.

Dr. LEVIN. But it has to go from elemental to inorganic before the methylation can occur.

Senator ALLARD. But my question is, is that the type of mercury that is being brought in from Asia, what form of mercury is that?

Dr. LEVIN. Because of its long-range transport, it is primarily the elemental form, but the atmospheric chemistry of mercury changes that progressively into the inorganic form, which is the form that readily deploys.

Senator ALLARD. Now, can the inorganic form be transferred into methyl-mercury?

Dr. LEVIN. Yes, sir. That is the form.

Senator ALLARD. So all those type of compounds get acted on by bacteria and then that is how that gets into the food chain.

Dr. RICE. The elemental form and the inorganic form are converted back and forth.

Senator ALLARD. I see.

Dr. RICE. So it does not make any difference whether it reaches the North American shores as elemental mercury or inorganic mer-

cury. Once it is deposited into the soil or the river, it is going to become inorganic mercury that then becomes available to be able to be turned into methyl-mercury.

Senator ALLARD. OK, thank you.

Now, here is the question, and I would like to have all of you respond to this. In your opinion, would a decrease in U.S. anthropogenic mercury emissions have an effect on global mercury levels? And part of the rest of the question is, apparently there is a high percentage of mercury present in the United States from outside our borders, so what effects can we expect from a decrease in our emissions? We have a couple of questions there and I would like to have all of you respond to those if you would.

Dr. RICE. There is no question that there is a global cycling of mercury. A lot of the mercury in the United States comes in from someplace else, comes in from the West, but some of it may have in fact originated in the United States originally. This stuff really does circle the globe. So just because it is coming in from the West does not mean it wasn't ours to start with.

Senator ALLARD. We do not know how much starts here.

Dr. RICE. No, we do not, and I am not a modeler so I really cannot speak to that. But what I do know is that there is local deposition. In other words, the mercury that is released from power plants in the Midwest ends up downwind. I just moved to Maine, and Maine is the so-called tailpipe for that local deposition, for that local emission. There is a percentage of it, and Dr. Levin can tell you what the percentage is better than I can, that is locally deposited. I think it is something like 30 percent.

Getting rid of those local sources would certainly at least help the Northeastern United States. Originally, the modeling, it was thought that this would take a long, long time. There are newer data now where small studies have actually been done that suggest that it might not be as grim as we originally thought; that these local changes can take place in a relatively shorter time, over the course of several years, rather than decades and decades as we originally may have feared.

Senator ALLARD. Dr. Myers, do you have a comment on that?

Dr. MYERS. It is outside of my area of expertise.

Senator ALLARD. Dr. Levin.

Dr. LEVIN. Dr. Rice is primarily correct on that. The deposition within the United States makes up about 30 percent of U.S. emissions. The rest of the emission go globally. Our modeling considered the fate of U.S. emissions and accounted for the amount that basically circles the globe and comes down after one trip around the world.

It is also correct that there is local deposition that in some cases may be significant near particular groupings of sources. I indicated that in my testimony, that although the average change in deposition for the scenario was 3 percent, there were some small areas where it was as much as 10 times that on a percentage basis.

So it calls for more detailed studies and particularly more looking at the science of tracking mercury found in fish back to its sources scientifically, that is, figuring out where it came from.

Senator ALLARD. Thank you, Mr. Chairman. I believe my time has expired.

Senator INHOFE. Yes, thank you.

Senator Carper.

Senator CARPER. Thank you, Mr. Chairman.

To our witnesses, again thank you for joining us. Thank you for your patience in bearing with us.

Dr. Rice, did I understand you to say you have concluded two decades of work at EPA?

Dr. RICE. Well actually most of it was not at EPA. I was at Health Canada for 22 years. I am American, but I graduated from the University of Rochester, got my Ph.D. from the University of Rochester so I have known Dr. Myers for many years. Then I went up there to work at Health Canada.

Senator CARPER. I see. Thank you for your service at EPA, and thank you all for real interesting testimony today.

Sometimes these are fairly technical issues. What is helpful for me as I listen to the comments of each of your testimonies and your responses to our questions is to look for threads of consensus; not to focus so much on where you disagree, but to find some areas where you agree. I would just ask each of you to take a minute or two and just to talk about some of the areas where you think you agree, and which might be helpful to us as we wrestle with whether to craft legislation, enact legislation along the lines that Senator Jeffords has introduced, I have introduced, or the President has proposed.

Can you help me with that? Dr. Levin, why don't you go first.

Dr. LEVIN. Thank you, Senator. We agree that mercury is a highly toxic compound. Its presence in the U.S. diet may in some instances cause concerns for development of children neurologically. We agree that there may be other effects that have to be looked for in terms of the health effects.

We also agree that the science of mercury is still emerging; that the linkage between health effects in particular areas, or for that matter in entire regions of the United States, and the sources of mercury is a critical question that would shape a wise course toward management decisionmaking. The work that I have been describing today is a step in doing that. The work that has been described by the other two witnesses today on health effects is a critical part of that linkage.

Bringing this source-receptor issue together with the health effects on a specific geographic basis and among specific populations within the United States is a key part in answering the management questions.

Senator CARPER. Thank you.

Dr. Myers, would you take a shot at my question please?

Dr. MYERS. I think we all agree that mercury is poisonous, every form. In high enough amounts, it is not only damaging to human health, but fatal generally. We all agree that it is worthwhile cleaning up the environment, I think. The question resolves at what level and at what cost. I think we all agree that these studies are extremely difficult to carry out and they are equally difficult to interpret because there are so many details to them. So it is so easy to end up with a bias either knowingly or unknowingly, generally I think unknowingly, that the interpretation of the details becomes incredibly important in these studies.

Senator CARPER. Thank you.

Dr. Rice.

Dr. RICE. I agree that we all know that methyl-mercury is toxic at high levels. There is absolutely no question about that. I agree with Dr. Myers that it is incredibly difficult to interpret these studies very often. They are very complex studies. There are a lot of variables, many of which we do not know. Epidemiology is an extremely blunt instrument. So that is why I think that it is important to look at the weight of evidence. There are a number of studies in humans that have documented effects of methyl-mercury at relatively low body burdens. In addition to that, there is a huge animal literature documenting effects and looking at the mechanisms of effects.

We do not know why one study may be positive, whereas another may be negative. So we really have to go with the evidence as a whole.

Senator CARPER. And maybe cite your most serious area of disagreement among you as panelists.

Dr. LEVIN. I would say disagreement probably rests in the question of the direction of research overall on the mercury issue, and how far that should continue.

Senator CARPER. Dr. Myers.

Dr. MYERS. I think the most serious area of disagreement is in the interpretation of the studies. We think that the Faroe Islands research is outstanding research. They have done a wonderful job. They have a great design. We are just not sure that they have been able to tease out from the mixture of chemicals present in whales a methyl-mercury component to it. That requires a lot of faith in their statistics and the details of the studies.

In the case of the New Zealand study, most people discounted the New Zealand study for many years. It was only when it was reanalyzed in the late 1990's that people began to start thinking of it in other terms. So I think our biggest disagreement is in the interpretation of it.

In addition, I think the weight of hundreds of small poorly done studies in difficult places such as the Amazon would never outweigh a really good study done looking at fish consumption.

Senator CARPER. Dr. Rice.

Dr. RICE. I guess everything that Gary Myers just said is my biggest point of disagreement. All of the smaller studies are not poorly done. Some of them are well done. The Faroe Islands study and the Seychelles study have been extensively reviewed. They are both considered to be very, very good studies.

The National Academy of Sciences looked at the issue of PCB coexposure very, very carefully and asked the investigators to go back and do a number of additional analyses. Their conclusion was that the effects seem to be independent of each other. These are both neurotoxicants. Although they both had effects in the study, the NAS conclusion was that they were independent.

Again, I think that we have go with a preponderance of evidence and not on just one study, no matter how well it has been done.

Senator CARPER. Mr. Chairman, I think this panel has been especially helpful to me. We thank you very, very much for your contributions today. Thank you.

Senator INHOFE. Thank you, Senator Carper.

Senator Clinton.

Senator CLINTON. Thank you, Mr. Chairman. I, too, want to thank the panel and welcome Dr. Myers from the University of Rochester, and Dr. Rice, your connection with Rochester, we will claim that as well.

I want to pick up where Dr. Rice just concluded. We have set up a system of evidence in our legal system that looks at the preponderance of evidence; that looks at a reasonable person standard. I share Dr. Rice's concern that we are not adequately responding to the evidence we already have, which I think the preponderance of it, certainly based on the review by the National Academy of Sciences, suggests that we have a problem with the transmission mostly in utero by mother to child that leads to neurological problems that in turn lead to poor school performance.

The 2000 report of the National Academy of Sciences found, I believe, that about 60,000 children might be born in the United States each year with this level of exposure that could affect school performance, but in your testimony you claim that more recent results from the CDC's National Health and Nutrition Examination Survey translate into over 300,000 newborns per year. Is that correct?

Dr. RICE. Yes. When the NAS did their analysis, the NHANES data was not available. The NHANES just started taking mercury blood and hair levels a couple of years ago, so those data have really become available since the NAS. They state that their 60,000 children was an estimate. It is actually about 320,000 children. Based on actual data that is representative of the U.S. population, it is above the EPA's reference dose.

Senator CLINTON. To me, this is truly alarming, that we have actual blood, hair sample, other kinds of physical examination which demonstrates that hundreds of thousands of our children are born each year potentially at risk for adverse affects on intelligence, memory, ability to pay attention, ability to use language and other skills.

Mr. Chairman, we are facing an increasing number of children in our school systems with learning disabilities. There are not any easy answers as to why the numbers of children with such learning disabilities has increased. Senator Jeffords has been a champion of making sure that all children are given an adequate education. In New York alone, we have 260,000 learning-disabled children. That is 50 percent of our special ed population. We spend $43 billion each year—$43 billion—on special ed programs for individuals with developmental disabilities between three and twenty-one.

Of course, not all special ed needs are the direct result of methyl-mercury exposure, but if it is demonstrably shown as we now have with evidence from the CDC's annual survey that we have levels of methyl-mercury in our children's bodies that is above what the EPA has determined to be healthy, and in fact some of us think the EPA standard is too low, but nevertheless if it meets that standard, then I would argue we have got to figure out how to address this environmental health challenge in a very short order.

I have been working with a number of colleagues to try to address the better data collection and environmental health tracking

that they need in the Individuals With Disabilities Act, and I think similarly on the scientific side with respect to better research and better analysis. But it is troubling to me that we are looking at a problem where the preponderance of the evidence I think is clear, where we know that there is a transmission, whether it is 60,000, 150,000, 300,000-plus children, and it needs some more effective response.

I wanted to ask you, Dr. Rice, now that you are in Maine, from the State perspective, how closely do you work with the State health department on environmental health issues? Do you exchange information with the State health department and even with the State education department about some of the work that you are doing?

Dr. RICE. I actually knew the State toxicologist for Maine quite well before I went up there, so I do interact with the health department. The methyl-mercury issue is very important to Maine. Maine has a very good program for trying to get rid of methyl-mercury from dental amalgams, from thermometers, from the kinds of things that can be controlled; to not put mercury in landfills because Maine understands that we are at the end of the pipeline for methyl-mercury deposition. Maine has a terrible problem with fish advisories. There are a lot of places where fish cannot be eaten in Maine because of the deposition of methyl-mercury.

So I do work closely with the folks over there, and in fact my way here was paid by the air office, the Maine air office because the State of Maine is so very concerned about this issue. Maine is rural and it is poor, and it cannot really absorb the consequences of these kinds of additional exposures on the health of the people of Maine.

Senator CLINTON. Similarly, new science is demonstrating that we need lower standards for lead, based on what we are now determining. A lot of that groundbreaking work was done at the University of Rochester about lead exposures and the impacts of lead exposure. We can take each of these chemicals or compounds piece by piece, but I think that certainly when it comes to mercury and lead and their impacts on children's development, it is not something I feel comfortable studying and waiting too much longer on, particularly because there are so many indirect costs. I know that Dr. Levin's work looked at some of the risks and cost-benefits, but people do not seem to factor in this special education population that has been growing.

Dr. RICE. If I may make a comment, I think your analogy is an apt one, and I think it is a very informative one. In 1985, there was a report to Congress on the cost-benefits of lead, of keeping lead out of gasoline, in fact. The benefits based on not only special education and things like lower birth weight with respect to lead, but also just the economic consequences of lowering the IQ of workers amounted to billions and billions of dollars a year in 1985 dollars or 1994 dollars. So as this effort goes forward in terms of figuring out how much it is going to cost to reduce mercury emissions, this other side of the equation, how much it is going to cost not to, needs to be kept very, very well in mind.

Senator CLINTON. Thank you, Dr. Rice.

Senator INHOFE. Thank you, Senator Clinton.

I thank the panel very much for their testimony.

Senator JEFFORDS. I had a couple more questions.

Senator INHOFE. Well, all right. It has to end at 12 o'clock. Go ahead.

Senator JEFFORDS. Dr. Levin, before setting a mercury max standard, would you agree that it makes sense for EPA to conduct a full modeling analysis of all available technology options and their emissions reduction potential, including the most stringent options?

Dr. LEVIN. Yes, Senator. I think it is important for EPA to carry out a parallel study as EPRI has done, and to make that study public, as we have as well. I am not aware yet that they have actually done any modeling of a max standard since there has been no official proposal of one yet.

Senator JEFFORDS. Dr. Myers, I believe your testimony is that the fish consumed with an average mercury content of 0.3 parts per million has about the same mercury concentration as commercial fish in the United States. What are the concentration in non-commercial fish?

Dr. MYERS. Are you talking about the United States or the Seychelles?

Senator JEFFORDS. In the United States.

Dr. MYERS. Well, all fish has some mercury in it. Most of the commercial fish in the United States, I understand, has less than ½ part per million, but some of the fish, I am not sure what the non-commercial ones are, but it can go up to over two or three parts per million in some freshwater fish.

Senator JEFFORDS. Dr. Rice and Dr. Myers, can you characterize the body burden of the pollutants like mercury in American children compared to the levels found in the Seychelles children?

Dr. MYERS. The average hair level in the mothers in Seychelles is 6.9 in the group we were studying. The average in the United States is less than one part per million. The average in Japan is somewhere around two parts per million.

Senator JEFFORDS. Dr. Rice, any comment?

Dr. RICE. No. That is correct, but I think it is important to understand that the NHANES data did identify some women, a very small percentage of women with higher hair mercury levels. I think it is important also to understand that the NHANES data are designed to be representative of the U.S. population as a whole, so that women who may eat more fish and may be at more risk for increased body burdens of methyl-mercury, such as immigrant populations or populations of people who are subsistence anglers and who eat inland fish. This is not captured. These populations are not captured by the NHANES data and I think that this needs to be kept in mind.

Senator JEFFORDS. I have some further questions I would like to submit.

Senator INHOFE. That would be perfectly appropriate. I appreciate it very much, and I appreciate the panel coming and also your patience from the long first session.

We are now adjourned.

[Whereupon, at 12 o'clock p.m. the committee was adjourned, to reconvene at the call of the chair.]

[Additional statements submitted for the record follow:]

STATEMENT OF HON. JON CORNYN, U.S. SENATOR FROM THE STATE OF TEXAS

Mr. Chairman, I commend you for holding this important hearing examining what is known about the science of climate change, mercury and the potential health effects of mercury emissions from power plants.

Given the timing of the energy debate on the Floor and this Committee's ongoing consideration of the Clear Skies Act, this is a very timely and important topic and I commend the Chairman for setting time aside to focus on the issue. I realize our focus today in regards to climate change is on the science, principally on temperature change. Two very different trains of thought are about to be presented to us today and I think this is positive and encourages a good, healthy debate. The question that this panel has to wrestle with is moving ahead with a greenhouse gas policy that may or may not be based on sound science. I am concerned about the costs in moving forward when there is a large body of science out there that says there isn't a problem.

To shift our focus just a bit, an issue of particular concern to me is the available technology to control greenhouse gas emissions, specifically CO_2. I am fairly certain that some of my colleagues agree with the line of thought about to be outlined by Dr. Mann, and this could very well lead this committee to a debate imposing mandatory controls on CO_2. If this turns out to be the case it is imperative that this Committee determine whether or not the technology is currently available to accomplish CO_2 reductions that are effective enough to solve the "problems" thought to be faced. I realize this is a topic for another hearing, but one that causes me concern.

In regards to mercury, in the 1990 amendments to the Clean Air Act, Congress specifically requested that EPA conduct an analysis of the health effects of mercury emissions from power plants and report back. EPA did conduct that study in 1997 and concluded that there was a "plausible link" between mercury emission and potential health effects, but was unable to quantify the link.

Six years have passed since EPA's 1997 study. Unfortunately, we still have not received any clarification from the EPA as to the magnitude of the health risks posed by power plant emissions, even though we are currently on the verge of spending billions of dollars to reduce those emissions.

I suspect that one of the reasons for this lack of information is that we are dealing with a global problem. Many people today may find it surprising to learn that most of the mercury that is deposited in the United States originates from outside our borders. In fact, for most of the country, over 60–80 percent of the mercury deposited in the United States comes from emission sources located in another country. Additionally, natural sources of mercury, such as forest fires and vegetation burning, account for over half of the world's mercury emissions.

What this means is that we have control over only a very small portion of total mercury emissions. Of the 5500 tons of mercury emitted globally, the U.S. accounts for only about 155 tons, or 3 percent of global emissions. U.S. power plant emissions which are estimated to be 48 tons per year, represent less than 1 percent of total global emissions. Given how small this fraction is, it is both reasonable and prudent to ask what impact controls on power plants will have on actual public health.

While EPA has unfortunately not provided us with any data on that question as of yet, Leonard Levin from the Electric Power Industry has. According to his very detailed analysis, control programs to reduce mercury emissions from power plants are likely to have less than a 1-percent impact on public exposure in this country. In fact, he estimates an impact of less than 0.3 percent. I do not know if this number is correct, but I think his very detailed analysis deserves comment from EPA, especially given that this was exactly the kind of information Congress sought in 1990 when it amended the Act.

I look forward to hearing Dr. Levin's testimony, as well as Dr. Rice's and Dr. Myers'. Your collective input is critical to this committee as we continue to debate the Clear Skies initiative.

I yield back the balance of my time.

STATEMENT OF DR. WILLIE SOON, HARVARD-SMITHSONIAN CENTER FOR ASTROPHYSICS

Distinguished Senators, panelists, and audience: My name is Willie Soon. I am an astrophysicist with the Harvard-Smithsonian Center for Astrophysics in Cambridge, Massachusetts. My training is in atmospheric and space physics and my sustained research interests for the past 10 years include changes in the Sun and their possible impact on climate.

This very rich area of scientific research, though still far from having definitive answers, has seen exciting and important progress from our increasing technical ability to measure, quantify, and interpret the changes in the Sun which could be linked to changes of the Earth's climate.

Today I focus on my latest research conclusions regarding climate change over roughly the last 1000 years, especially the geographical pattern of those changes. My scientific study is only possible because of the careful research produced by nearly one thousand scientists around the world. Their expertise covers a very wide range, including physical, chemical, biological, and geological sciences.

Together with several colleagues whose names are listed in the two scientific papers that I am submitting today for the record of this testimony, we have synthesized the results from several hundred studies of proxy records of climate, including much new work that has appeared in the scientific literature in the last 5 to 10 years.

Climate proxies are indirect climate sensors based on information from tree rings, ice and seafloor sediment cores, corals, glaciers and other natural evidence. They also include important cultural and documentary records.

It is important to recognize that these climate proxies are not temperature readings, but some proxies may be calibrated to give temperature changes. One example is the measurement of the flow of heat in boreholes drilled through rocks or ice, yielding century-scale temperature changes over several millennia. On the other hand, some proxies are sensitive to local rainfall as well as temperature, as in the case of annual tree growth in the southwest United States. Any given proxy may respond to temperature differently from other proxies, depending on, for instance, the type of proxy, location, or season.

For all those reasons, it remains a big challenge to produce an accurate global temperature record over the past 1000 years from the diverse set of climate proxies.

But within the limits and lessons learned from our research papers, we can offer three conclusions:

First, local and regional, rather than "global", changes are the most relevant and practical measure of climate change and impact. This is because truly global averages rarely are available from the distant past, before modern satellite measurements, and because such averages can hide the significant changes that can occur over large parts of the Earth.

Second, on a location by location basis, there was a widespread Medieval Warm Period between approximately 800 and 1300 A.D. This Medieval Warm Period was followed by a widespread colder period, called the Little Ice Age, that lasted from approximately 1300 to 1900 A.D.

Third, there is no convincing evidence from each of the individual climate proxies to suggest that higher temperatures occurred in the 20th century than in the Medieval Warm Period. Nor is there any convincing evidence to suggest that either the rate of increase or the duration of warming during the 20th century were greater than in the Medieval Warm Period.

The fact that local and regional climate has been varying with significant swings in amplitude over many locations provides important challenges for computer simulation of climate. The full models that explore the Earth region by region can test for the natural patterns of change over the last 1,000 years through the use of the climate proxies we just discussed. In that way, the effects of human-caused climate change can be weighed against observed natural variability in the climate system. Having computer simulations reproduce past climate, which has been influenced predominantly by natural factors, is key to making an accurate forecast that includes all potential human-made warming and cooling effects.

Further research could yield a deeper, quantitative improvement to our knowledge of local and regional climate variability during the past 1000 years. As we could be inspired by Mr. Thomas Jefferson who remarked:

> "It is a common opinion that the climates of the several states of our union have undergone a sensible change since the dates of their first settlements; that the degrees of both cold & heat are moderated. The same opinion prevails as to Europe; if facts gleaned from history give reasons to believe that, since the times of Augustus Caesar, the climate of Italy, for example, has changed regularly at the rate of 1 [degree] of Fahrenheit's thermometer for every century. May we not hope that the methods invented in latter times for measuring with accuracy the degrees of heat and cold, and the observations which have been & will be made and preserved, will at length ascertain this curious fact in physical history?"—Marginal notes from Thomas Jefferson's Monticello Weather Diary (January 1, 1810 to December 31, 1816).

I strongly believe that the time for research in paleoclimatology to fulfill this important role is now.

RECONSTRUCTING CLIMATIC AND ENVIRONMENTAL CHANGES OF THE PAST 1000 YEARS: A REAPPRAISAL

Willie Soon (wsoon@cfa.harvard.edu) and Sallie Baliunas
Harvard-Smithsonian Center for Astrophysics, Cambridge, Massachusetts and Mount Wilson Observatory, Mount Wilson, California
Craig Idso and Sherwood Idso
Center for the Study of Carbon Dioxide and Global Change, Tempe, Arizona
David R. Legates
Center for Climatic Research, University of Delaware, Newark, Delaware

ABSTRACT
The 1000-year climatic and environmental history of the Earth contained in various proxy records is examined. As indicators, the proxies duly represent or record aspects of local climate. Questions on the relevance and validity of the locality paradigm for climatological research become sharper as studies of climatic changes on timescales of 50–100 years or longer are pursued. This is because thermal and dynamical constraints imposed by local geography become increasingly important as the air-sea-land interaction and coupling timescales increase. Because the nature of the various proxy climate indicators are so different, the results cannot be combined into a simple hemispheric or global quantitative composite. However, considered as an ensemble of individual observations, an assemblage of the local representations of climate establishes the reality of both the Little Ice Age and the Medieval Warm Period as climatic anomalies with world-wide imprints, extending earlier results by Bryson et al. (1963), Lamb (1965), and numerous other research efforts. Furthermore, these individual proxies are used to determine whether the 20th century is the warmest century of the 2nd Millennium at a variety of globally dispersed locations. Many records reveal that the 20th century is likely *not* the warmest nor a uniquely extreme climatic period of the last millennium, although it is clear that human activity has significantly impacted some local environments.

KEY WORDS: Paleoclimate proxies; Climate change; Environmental change; Little Ice Age; Medieval Warm Period.

1. INTRODUCTION

Are the Little Ice Age and Medieval Warm Period widespread climatic anomalies? Nearly four decades ago, H. H. Lamb (1965, pp. 14–15) wrote, "[M]ultifarious evidence of a meteorological nature from historical records, as well as archaeological, botanical and glaciological evidence in various parts of the world from the Arctic to

New Zealand... has been found to suggest a warmer epoch lasting several centuries between about A.D. 900 or 1000 and about 1200 or 1300... Both the "Little Optimum" in the early Middle Ages and the cold epochs [i.e., "Little Ice Age"], now known to have reached its culminating stages between 1550 and 1700, can today be substantiated by enough data to repay meteorological investigation... It is high time therefore to marshal the climatic evidence and attempt a quantitative evidence." In response to Lamb's call to action, research on large-scale patterns of climate change continued with vigour.

Thirty-three years later, however, Jones et al. (1998) tentatively concluded that "[w]hile the 'Little Ice Age' cooling (with the seventeenth century being more severe over Eurasia and the nineteenth century more severe over North America) is clearly evident ... we can only concur... that there is little evidence for the 'Medieval Warm Period'... although the fact that we have only four series before 1400 and the timescale limitations described earlier [i.e., not resolving timescales of multidecades to century with tree ring proxies used in their study] caution against dismissing the feature."

Overpeck et al. (1997) had previously commented that "[t]he annually dated record of Arctic climate variability encompassing the last 1000 years has less spatial coverage than does the multiproxy record of the last 400 years. Sediment, ice core, historical, and tree ring data for this earlier period indicate that although Arctic summers of the 20th century were generally the warmest of the last 400 years, they may not be the warmest of the last millennium[1]... The few time series of climate change spanning the last millennium also suggest that the Arctic was not anomalously warm throughout the so-called Medieval Warm Period of the 9th to 14th centuries." Nevertheless, the updated composite tree-ring summer temperature curve in Figure 1 of Briffa (2000) shows clear evidence of an anomalously warm interval from about 950 to 1100 A.D. in the northern high-latitude zone, which coincides with the Medieval Warm Period discussed here. Also, an early warm period appears prominently in the averaged tree ring chronologies carefully selected and processed from 14 sites spread over 30–70°N latitude (Esper et al. 2002a).

These results are but a few of the many that have become available since Lamb's pioneering analysis. Given advancements in retrieval of information from climate proxies, as well as their extensive surface coverage, we review the accumulated evidence on climatic anomalies over the last 1000 years. We also recommend the study of Ogilvie and Jónsson (2001), which provides the most authoritative, up-to-date discussion of the historical development of the long-standing debates over the climatic nature of the Medieval Warm Period and Little Ice Age, especially in the regions surrounding the North Atlantic, including Iceland.

---

1 When considering the possible link of early 20th century warming to the rise in atmospheric CO_2 concentration, it should be noted that the Arctic-wide temperatures of Overpeck et al. began rising in the mid-19th century and peaked around 1940–1960, when the increase in the air's CO_2 content was less than 20–30% of the cumulative CO_2 increase to date; see Etheridge et al. (1996) for the preindustrial level of CO_2.

2. WORKING DEFINITIONS

First, working definitions of the Medieval Warm Period and Little Ice Age must be established in order to assess the various climate proxy records. For example, Grove (2001a) captures the difficulty in deciphering the nature of the Medieval Warm Period and Little Ice Age: "The term "Little Ice Age" does not refer directly to climate but to the most recent period during which glaciers extended globally and remained enlarged, while their fronts fluctuated about more forward positions... The term Medieval Warm Period has been the subject of considerable controversy. Its nature and even its existence has been queried,... as has that of the Little Ice Age... They were not periods of unbroken cold and warmth respectively. Climate varied on small scales both spatially and temporally, as it has also in the twentieth century. Nevertheless, climatic conditions were such during the Little Ice Age that mass balances were sufficiently predominant for the glaciers to remain enlarged, although their fronts oscillated. Similarly during the Medieval Warm Period climatic conditions caused mass balances to be negative, and volumes of glaciers to be reduced, so that they retracted substantially, though their fronts no doubt fluctuated, as they have been observed to do during the warming of the twentieth century."

Lamb (1982, 1997a), also mindful of the complex nature of weather and climate, noted that: "within the last thousand years, the development of what has been reasonably called the Little Ice Age seems to have affected the whole Earth, as has the twentieth-century recovery from it; but when the ice on the Arctic seas extended farthest south, particularly in the Atlantic sector, all the climatic zones seem to have shifted south, including the storm activity of the Southern Ocean and the Antarctic fringe. This apparently broke up much of Antarctic sea ice, enabling Captain Cook in the 1770s and Weddell in 1823 to sail further south than ships have usually been able to reach in this century.[2] The southward extension of open water would presumably result in some mildening of the regime not only over the ocean but some way into the interior of Antarctica, and this just when the world in general north of 40°S was experiencing a notably cold regime. Amongst the evidence which builds up this picture, at that time the winter rains failed to reach so far north over Chile. And radiocarbon dating of abandoned penguin rookeries on the Antarctic coast near 77.5°S, in the southernmost part of the Ross Sea, suggests that there were periods of milder climate there about AD 1250–1450 and 1670–1840. These periods include the sharpest phases of development of the Little Ice Age climate in the northern hemisphere." (page 39 of Lamb 1997a)

What are the regional and global patterns of climatic change over the last 1000 years? Accurate answers to these questions are important, both as benchmarks for the 20th century global average warming exhibited by surface thermometer records and as

2 See e.g., the evidence (Hendy et al. 2002) for relative warmth in the reconstructed coral-isotopic sea surface temperature throughout most of the 18th and 19th centuries at the central Great Barrier Reef, Australia. It should be noted, however, that this single proxy result does not imply uniform warmth throughout the whole south Pacific, south Atlantic and Indian Oceans. For historical accounts of sea ice conditions and harsh weather extremes during Cook's second voyage, see e.g., Forster (2000).

physical constraints for theories or mechanisms of climate change on timescales of decades to centuries.

To make progress towards this understanding, we address three questions of many individual climate proxies that differ too widely to be quantitatively averaged or compared:

(1) Is there an objectively discernible climatic anomaly occurring during the Little Ice Age, defined as 1300–1900 A.D.? This broad period in our definition derives from historical sea-ice, glaciological and geomorphological studies synthesized in Grove (2001a, 2001b) and Ogilvie and Jónsson (2001).

(2) Is there an objectively discernible climatic anomaly occurring during the Medieval Warm Period, defined as 800–1300 A.D.? This definition is motivated by Pfister et al. (1998) and Broecker (2001) and is slightly modified from Lamb's original study (1965).

(3) Is there an objectively discernible climatic anomaly occurring within the 20th century that may validly be considered the most extreme (i.e., the warmest) period in the record? An important consideration in answering this question is to distinguish the case in which the 20th century warming began early in the century versus after the 1970s, as recorded by surface thermometers. This criterion is necessary in order to judge the influence of 20th century warming by anthropogenic forcing inputs such as increased atmospheric carbon dioxide content.

Anomaly, in our context, is simply defined as a period of 50 or more years of sustained warmth, wetness, or dryness within the Medieval Warm Period, or a 50-year or longer period of cold, dryness, or wetness during the Little Ice Age. Definition of the 20th century anomaly is more difficult to establish. The 20th century surface instrumental temperature record contains three distinct, multidecadal trends: early-century warming, mid-century cooling and late-century warming. But that knowledge comes from instrumental thermometry with its high time resolution and other biases that preclude a direct comparison with the proxies (proxies have their own biases). Hence, a further aspect of our research goal is to compare the 20th century objectively with more extended temperature and precipitation histories than are available from instrumental records. Given the biases of each proxy, question 3 was answered by asking if, within each proxy record, there were an earlier (pre-20th century) 50-year interval warmer (or more extreme, in the case of precipitation) than any 50-year average within the 20th century.

The third question differs from the first two. Question 3 seeks a 50-year anomaly within the 20th century compared to any other anomaly throughout the period of a proxy record while Questions 1 and 2 search for 50-year anomalies within the previously suggested 500-year and 600-year intervals of the Medieval Warm Period and Little Ice Age, respectively. But note that in the case of the third question, we treat the definition of a 50-year or more period of sustained anomaly in the 20th century no differently from that of any prior century. Thus, if a sustained warm anomaly were identified during the Medieval Warm Period and appeared warmer than an anomaly found in the 20th century, then we would assign 'No' to question 3. Similarly, a proxy record may show, for example, both that the 20th century anomaly is the most extreme

(warmest) and that the Medieval Warm Period exists. In answering the third question, the existence of the Medieval Warm Period or Little Ice Age is not considered as they are assessed independently in answering the first two questions.

We begin with the framework of past researchers; namely, the suggested existence of the Medieval Warm Period and Little Ice Age. Our goal is thus to deduce the geographical nature of climatic and environmental conditions during these periods. Distinguishing the 20th century as a separate period is a result of the interest in the role of human activity on Earth's climate and environment.

Another important consideration is that temporary regional cooling may have occurred on decadal, but not on multidecadal, timescales during the Medieval Warm Period, and that occasional, short-lived regional warming may have occurred during the Little Ice Age, as indicated by J. Grove (2001a, 2001b). Use of the terms Medieval Warm Period and Little Ice Age should suggest persistent, but not necessarily constant, warming or cooling, respectively, over broad areas (see Stine 1998, Grove 2001a, 2001b, Luckman 2000, Ogilvie and Jónsson 2001; Esper et al. 2002a). As suggested by Stine (1998), therefore, more appropriate terms may be the 'Little Ice-Age Climatic Anomaly' and the 'Medieval Climatic Anomaly'. Also note that the definitions of discernible, persistent climate anomalies for the Little Ice Age and Medieval Warm Period include not only distinct changes in the climatic mean but also changes in multidecadal variations (Ogilvie and Jónsson 2001). In the context of daily and regional spatial scale variability, it is important to recognize that the relationship between multidecadal mean temperature and its daily variability may undergo significant non-stationary changes (see Knappenberger et al. 2001, who document those specific time-dependent changes in temperature variability across the United States). Also, from a combination of field evidence and modelling based on an understanding derived from synoptic climatology, Bryson and Bryson (1997) demonstrated how local and regional factors (for horizontal spatial distances as small as 100 km) have produced significantly different precipitation histories for two Near East stations (e.g., Jerusalem and Kameshli, Syria) and for two stations in the Cascade Range of Oregon (e.g., mountainous versus coastal-like microclimate locations).

Our classification of a widespread anomaly based on multidecadal persistence at many locales rests on good precedent. For example, the modern globally averaged surface warming inferred from thermometer readings includes large-scale cooling trends over both the Greenland/Labrador Sea area and the eastern region of the United States (e.g., 30–45°N; 80–110°W; see Hansen et al. 1999, Robinson et al. 2002) or the Antarctic continent (e.g., Doran et al. 2002) in the last 30–50 years. Another example is the relative warmth during the Little Ice Age and relative coolness during the Medieval Warm Period seen in the borehole record of reconstructed temperature at Taylor Dome, Antarctica (77.8°S, 158.72°E, elevation 2374 meters), as compared to results from Greenland's borehole (see Clow and Waddington 1999), which do not show those features.

Assessing and confirming the global extent of the Little Ice Age and the Medieval Warm Period is premature because proxy data are geographically sparse and either one or both phenomena could be multi-phased events acting under distinct local and regional constraints and modes. Bradley and Jones (1993) and Hughes and Diaz

(1994) initiated and championed the position for a non-global concomitance of the phenomena (but consider pp. 51–54 of Grove 1996 for an important clarification regarding this discourse, in light of the evidence for the Little Ice Age from glacial geology). However, in the traditionally data-rich areas of Western Europe and the Northern Atlantic, including Iceland and Greenland, both the Little Ice Age and Medieval Warm Period were distinct climate anomalies (see e.g., Pfister et al. 1998; Grove 2001a; Ogilvie and Jónsson 2001) and no objective proof discredits the existence of those phenomena in other regions. Thus, consistent with other researchers (e.g., Lamb 1965; Porter 1986; Grove 1996; Kreutz et al. 1997), we assume that both the Little Ice Age and the Medieval Warm Period may be globally distributed and teleconnected events that need not necessarily imply an extended period of global cooling or warming that persisted uniformly throughout the defined durations. The terms Medieval Warm Period and Little Ice Age still appear practical and viable, especially when considering their extension to past and future climatic events that are 'similar or equivalent' in physical scope (e.g., Bond et al. 1997, 1999; Khim et al. 2002; Stott 2002; Stott et al. 2002).[3]

Even with limited knowledge of the diverse range of local climatic behaviour, the Medieval Warm Period and the Little Ice Age are not expected to be spatially homogeneous or temporally synchronous. The beginning and end dates of these climate anomalies require a better understanding (e.g., for the Little Ice Age see Porter 1981, 1986; Kreutz et al. 1997; Kaser 1999; Grove 2001a, 2001b; Luckman 2000; Schuster et al. 2000; Winkler 2000; Ogilvie and Jónsson 2001; Hendy et al. 2002; Mayewski and White 2002; Qian and Zhu 2002; Paulsen et al. 2003). Also, the imprecision of the timing of both events contributes, in part, to confusion about the phenomena. For example, Ogilvie and Farmer (1997) commented that Lamb's suggestion of a Medieval Warm Period may not be supported by documentary data even for England, because Ogilvie and Farmer's extensive and careful research using an historical dataset showed that England suffered relatively cold winters from 1260 to 1360 A.D. However, that period is near our transition between the Medieval Warm Period and Little Ice Age, so this fact does not strongly challenge our working definition and research. Collected evidence, especially that based on glacier activity, points to both a diffuse beginning and end of the Medieval Warm Period, while the Little Ice Age interval seems to have had a diffuse beginning but a more abrupt end. Based on Na^+ concentration records from annually dated ice cores from central Greenland and Siple Dome, West Antarctica, Kreutz et al. (1997) showed that the onset of Little Ice Age conditions around both poles seems to be abrupt and near-synchronous, starting about 1400 A.D. Although the notion of a Medieval Warm Period or a Little Ice Age with sharply defined transitions may be a convenient one, it is probably a non-physical construct, because regional differences in the timing of both phenomena could be quite large. As hinted by Grove (2001a), a similarly

3 Note that we refer to distinct cold and warm phases together with corresponding expressions of glacial changes but without the acceptance or refutation of the important "sharp spectral line" controversy related to climate system variability on millennial timescale discussed between Wunsch (2000, 2001) and Meeker et al. (2001).

inhomogeneous climate pattern also can be identified in the 20th century warm interval.

We offer an overview of a multitude of research results within our idealized framework to address our three questions about the existence of climate anomalies at individual locations. Climate indicators considered include information obtained from documentary and cultural sources, ice cores, glaciers, boreholes, speleothems, tree-growth limits, lake fossils, mammalian fauna, coral and tree-ring growth, peat cellulose, pollen, phenological data, and seafloor sediments. In its own way, each proxy provides a unique view of climate variability in terms of its relative sensitivity to the planet's thermal and hydrological fields, as well as nonclimatic factors. We rely on individual researchers for their best judgements in identifying the most significant climatic signals in their studies. Thus, our three questions are addressed within the context of local or regional sensitivity of the proxies to relevant climatic variables, including air temperature, sea surface temperature, precipitation and any combination of large-scale patterns of pressure, wind and oceanic circulation.

3. UNCERTAINTIES IN INFERRING CLIMATE FROM PROXIES

The accuracy of climate reconstruction from proxies, including the awareness of anthropogenic interventions that could pose serious problems for a qualitative and quantitative paleoclimatology, was discussed by several researchers, including Bryson (1985) and Idso (1989). Temperature changes inferred for the Medieval Warm Period and Little Ice Age Climatic Anomalies are generally accepted to be no more than 1 to 2°C when averaged over hemispheric or global spatial scales and over decades to a century. Broecker (2001) deduced that only the results from mountain snowline and borehole thermometry are precise to within 0.5°C in revealing changes on centennial timescales. But the quantification of errors is complex, and both Bradley et al. (2001) and Esper et al. (2002a) have challenged Broecker's statement. In addition, Jones et al. (1998) have provided an excellent review on the quantitative and qualitative limitations of paleoclimatology; and Henderson (2002) provides more detailed cautions on the limitations of various climate proxies, as well as an excellent overview on potential new oceanic proxies. Others, such as Ingram et al. (1978) and Ogilvie and Farmer (1997), have cautioned about quantitative interpretations of climatic results based on historical documentation.

In our independent survey of the literature, we have observed three distinct types of warnings (see Bryson 1985; Clow 1992; Graybill and Idso 1993; Huang et al. 1996; Briffa et al. 1998; Cowling and Skyes 1999; Schleser et al. 1999; Evans et al. 2000; Schmutz et al 2000; Aykroyd et al. 2001; Ogilvie and Jónsson 2001; Zorita et al. 2003): (i) lack of time-scale resolution for the longest-term component of climate signals; for example, in tree ring and coral records, or the loss of short-term climate information in borehole temperature reconstructions; (ii) nonlinearities (i.e., related to age, threshold, discontinuous or insufficient sampling, saturated response, limited dynamic range of proxy, etc.) of biological, chemical and physical transfer functions necessary for temperature reconstruction; and (iii) time dependence or nonstationarity of the climate-proxy calibration relations.

Estimates of ground temperature trends from borehole data can be complicated by

non-climatic factors associated with changes in patterns of land-use and land cover over time (Lewis and Wang 1998; Skinner and Majorowicz 1999). In general, climate proxies from floral and faunal fossils in lake and bog sediments are only sensitive enough to resolve change to within ±1.3°–1.8°C (e.g., Lotter et al. 2000). Isotope-coral proxies lack the climate-sensitivity resolution and the continuous length of coral growth to address millennial climatic change. Jones et al. (1998) demonstrated that both coral- and ice-core based reconstructions performed more poorly than tree-ring records when calibrated against thermometer data since 1880 A.D. In contrast, the tree ring proxy that has the best annual time-resolution is limited by the loss of information on centennial components of climate change (see further discussion in section 5.1).

Amplitudes of large-scale surface temperature change derived from tree-ring proxies can be substantially underestimated – by a factor of two to three as compared to results from borehole thermometry (Huang et al. 2000; Harris and Chapman 2001; Huang and Pollack 2002). It seems surprising that the amplitude of climate variability broadly resolved by borehole reconstruction on timescales of at least 50 to 100 years is larger than the high time resolution results from tree-ring proxies, rather than lower, because short-term climate fluctuations are smoothed out by the geothermal heat-flow that acts as a low-pass filter.[4] Differing amplitudes resulting from borehole and tree ring climate proxies suggest that longer time scale (multidecadal and century) variability is more faithfully captured by borehole results, while the same information can be irretrievably lost in tree ring records (e.g., Cook et al. 1995; Briffa et al. 2001; Collins et al. 2002; Esper et al. 2002a) because of the standardization procedure (to remove nonclimatic biases related to changing tree size, age and tree-stand dynamics).[5] This is why Jones et al. (1998) commented that although one may be confident of comparing year-to-year and decade-to-decade (limited to periods shorter than 20–30 years) variability, which should be more sensitively imprinted in tree ring records, it requires "considerable faith" to compare, for example, the climate of the twelfth and twentieth centuries from tree ring proxies. To date, the goal of combining information from borehole and tree-ring proxies, or even between borehole and thermometer data, to arrive at a true proxy that simultaneously resolves timescales of years to centuries, has not been realized.

4 There are exceptions in careful tree-ring studies like those of Esper et al. (2002a), which employ new databases and strategies that are optimized to capture longer timescale variability; see further discussion in section 5.1.

5 Cook et al. (1995) cautioned that such standard detrending methodology, typically done by fitting the biology-related trend with a modified negative exponential curve or line of zero or negative slope (for the purpose of tree-ring dating, a high-pass filter is typically used), implicitly ignored the fact that a climatic signal could involve timescales exceeding the length of any individual segments of the full tree-ring chronology and even the maximum lifespans of tree species studied. In essence, the standardization process would indiscriminately remove both the biological trends and the variabilities driven by any slow changes in climate. There is thus a maximum climatic timescale that is resolvable by tree-ring proxy which in turn is related to the lengths of the individual tree-ring series. Such a general problem in the development of millennia-long tree-ring chronology has been labelled the 'segment length curse' by Cook et al. (1995).

Despite complicating factors such as the mismatch of climate sensitivities among proxies, Beltrami et al. (1995) and Harris and Chapman (2001) have begun to address the issue. Beltrami and Taylor (1995) also have successfully calibrated a 2000-year oxygen isotope record from an ice core near Agassiz with the help of borehole temperature-depth data near Neil for the Canadian Arctic. Their procedure avoids reliance on the circular approach of tuning a composite record from two different efforts to forcibly agree with each other (as illustrated in Figure 3 of Mann 2001a). Careful research such as that of Beltrami, Harris, Chapman and their colleagues may solve the difficulty of interpreting climate signals that degrade with borehole depth or time, which has led to the false impression that reconstructed temperatures contain a significantly smaller variability in the distant past than at present.

As long as no testable climate theory capably incorporates all relevant thermal, hydrological, geological, chemical, biological and other environmental responses during the Little Ice Age and Medieval Warm Period, local climatic proxies remain the most powerful benchmarks or measures of reality. Adopting the simplest classification that provides an objective answer to each of the three questions posed yields three advantages:

(a) the classification relies on local representations of climatic change, which are prerequisites for the construction of regional and global patterns of climate signals.

(b) the current application of mathematical decomposition techniques, such as empirical orthogonal functions (EOFs), to the world-wide reconstruction of a 1000-year temperature history is strongly limited by both the inhomogeneous spatio-temporal sampling gaps in proxy records (e.g., Evans et al. 2002; Schneider 2001) and the very short length of surface thermometer record available for calibration-verification purposes (discussed further in sections 5.1 and 5.4). The classification of proxies in this study is complementary to the mathematical decomposition processes but avoids some of their difficulties (albeit at the expense of quantitative results).

(c) the different sensitivities of proxies to climate variables and the time-dependence of the proxy-climate correlation (e.g., Briffa et al. 1998; Shabalova and Weber 1998; Schmutz et al. 2000; Aykroyd et al. 2001) require careful calibration and verification on a location by location basis; the classification's emphasis on local results avoids the difficulty of intercomparing disparate proxies. In other words, we are avoiding the over-emphasis on quantitative synthesis while keeping in mind that even for the same location, different proxies may yield different climate expressions simply because of their different sensitivities to local climatic variables (e.g., Fritz et al. 2000; Betancourt et al. 2002).

The combination of these three advantages suggests that a compact mathematical representation of individual proxy variations, without full understanding of proxy-climate calibration relations, may yield overconfident results. For example, Ambaum et al. (2001) and Dommenget and Latif (2002) studied the physical nature of the North Atlantic Oscillation (NAO) air pressure pattern derived from local one-point correlation analysis, in contrast to the pattern derived from an EOF analysis. A direct

association to local centres of action in the Euro-Atlantic region, which establishes the phenomenon of NAO, could not be found in the EOF representation. Dommenget and Latif (2002) showed more examples of mismatches in the dominant modes of variability deduced from EOF and one-point correlation analyses over the tropical Atlantic and tropical Indian Ocean (see the illuminating synthetic example shown in that paper). Tomita et al. (2001) cautioned that the family of EOF analyses may artificially over-emphasize standing or propagating features over a region with large variance. They specifically pointed out that EOF analysis tends "to extract a widespread variability over the tropics... and may fail to identify smaller but equally significant signals of DSV [Decadal Scale Variability] in coastal regions and/or the extratropical oceanic frontal zones." In other words, EOFs may be convenient measures for characterizing (and then deducing information for data-poor regions back to 1000 A.D., e.g., Mann et al. 1998, 1999, 2000a) dominant spatial-temporal components of climate variability, but they do not guarantee physical meaning and, hence, physical reality.

Our study has the disadvantage of being non-quantitative and very 'low-tech'. Thus, our assessment falls short of Lamb's (1965) original call for quantitative answers. In addition, by enforcing our simple rule of employing a local or regional perspective, we avoid difficult questions about the spatio-temporal coupling of observed changes over various regions, as well as questions about specific large-scale patterns that may be responsible for those climatic anomalies (see e.g., a particular framework/viewpoint, with the emphasis on the primary role of the "Mobile Polar Highs" which "organize migratory units of circulation in the troposphere low levels" in Leroux 1993 and further insights on the role played by stratospheric polar vortex in D. Thompson et al. 2002). An early effort to study the interlinkage of widely separated proxies, for example, between marine sediments at Palmer Deep, Antarctica and atmospheric signals in Greenland ice cores, was reported by Domack and Mayewski (1999). However, many chronologies produced by radiocarbon dating have an accuracy that is too limited to allow for a reliable discussion on the timing of events from different areas (see Stine 1998; Domack and Mayewski 1999; Khim et al. 2002). The difficult effort of areal weighting of different proxy records was attempted, for example, for the Arctic region by Overpeck et al. (1997), the Northern Hemisphere extratropics by Esper et al. (2002a), and both the Northern Hemisphere and global domains by Mann et al. (1998, 1999, 2000a). However, Briffa et al. (2001) criticized the lack of discussion of uncertainties in most (except in Mann et al. and Esper et al.) of these reconstruction efforts. For example, the composite series in the Overpeck et al. (1997) reconstruction is not even calibrated directly to instrumental data.

Our different approach to climate proxies may help to clarify the existing confusion concerning the non-local EOF-based reconstruction of global temperature by Mann et al., which often seems to differ from many local temperature proxy indicators (e.g., Bradley et al. 2000; but see also the reply to that commentary by Barnett and Jones 2000). We differ from Bradley et al.'s (2000) conclusion in that we believe that the spatial and temporal sampling of the available proxy network is not adequate for a robust capturing of the spatial pattern of changes on timescales greater than several

decades (see discussion in section 5.1).

4. RESULTS

Table 1 lists the world-wide proxy climate records we have collected and studied. In order to reduce the number of entries, the list is restricted to records that contain either direct information about the three specific questions posed earlier or that provide a continuous time series for at least 400–500 years. In addition, information was excluded from research in progress (e.g., the record of sea surface conditions around the Santa Barbara Basin with 25-year data resolution by J. Kennett, private communication 2001), as well as early results that may need other independent reexamination (e.g., Pisias's [1978] analysis of marine microfossils, *Radiolaria*, from varve sediments of the Santa Barbara basin). For the majority of cases, we strictly followed the individual researchers' statement about their paleoclimatic reconstruction efforts; but a few cases exist where our own judgements were imposed, based on requirements of consistency.

Table 1 (commences on page 290) and Figures 1–3 (over page) summarize the answers to the three questions we posed. For questions 1 and 2, we find the answer to be 'Yes' when the proxy record shows a period of 50 years or longer of cooling, dryness or wetness during the Little Ice Age and a period of 50 years or longer of warming, wetness or dryness during the Medieval Warm Period. A dash indicates that either the expert opinion or its logical extension is inconclusive, provides no information, or does not cover the period in question. A 'Yes?' or 'No?' answer means that the original expert opinion made a claim that does not match our criteria. For example, the interval of warmth during the Medieval Warm Period may be too short by our definition to be a 'Yes' and so we provide a 'Yes?' Finally, in several cases in the 20th century, a 'Yes*' designation was assigned for the answer to question 3 when the 20th century warming in the proxy records first occurred early in the century (ca. 1920–1950), when the air's anthropogenic CO_2 content was still cumulatively small. Our choice for the lower limit of 50 years for the recognition of a climatic anomaly is not entirely arbitrary as it grossly represents the longer periods unresolved by Mann et al.'s (1998, 1999, 2000a) multiregression analyses.

Our figures show the results of Table 1 for the Little Ice Age (Figure 1), Medieval Warm Period (Figure 2) and the nature of 20th-century's change (Figure 3). These figures graphically emphasize the general lack of climatic information extending back to the Medieval Warm Period for at least seven geographical zones: the Australian and Indian continents, South East Asian archipelago, large parts of Eastern Europe/Russia, the Middle Eastern deserts, the tropical African and South American lowlands (although the large number of available borehole-heat flow measurements in Australia seems adequate for the reconstruction of ground temperatures back to Medieval times; e.g., Huang et al. 2000). Therefore, our conclusions are provisional.

Figure 1 indicates the Little Ice Age exists as a distinguishable climatic anomaly in all regions of the world that have been assessed. Only two records – tree-ring growth from western Tasmania and isotopic measurements from ice cores at Siple Dome, Antarctica – do not exhibit any persistent or unusual climatic change over this period (although the western Tasmanian reconstruction contains an exceptionally cold decade

centred around 1900 A.D.; Cook et al. 2000).[6]

Figure 2 shows the Medieval Warm Period with only two unambiguous negative results. The Himalayan ice core result of Thompson et al. (2000) seems unambiguous, but the tree-ring proxy data from Lenca, Southern Chile (Lara and Villalba 1993) is countered by nearby evidence of the Medieval Warm Period (Villalba 1990; Villalba 1994).

Figure 3 shows that most of the proxy records do not suggest the 20th century to be the warmest or the most extreme in its local representations, which seems surprising until one realizes the more limited and contrary view was drawn primarily from

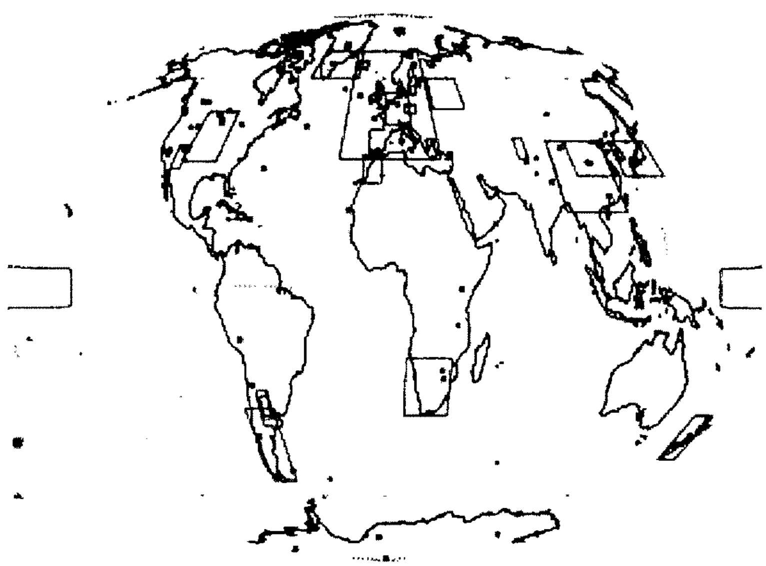


Figure 1. Geographical distribution of local answers to the following question: Is there an objectively discernible climatic anomaly during the Little Ice Age interval (1300–1900 A.D.) in this proxy record? 'Yes' is indicated by red filled-squares or unfilled boxes, 'No' is indicated by green filled-circles and 'Yes? or No?' (undecided) is shown with blue filled-triangles.

6 Cook et al. (2000) attempted to show that the warm-season climate retrieved from their Mount Read- Lake Johnston Huon pine tree-ring data is associated with the inter-decadal (>10-year) component of sea surface temperature in the southern Indian Ocean and southwestern sector of the Pacific ocean. However, it may not be appropriate to calibrate ring-growth occurring in the heavy rainfall region of the Western Tasmanian climatic zone by using the three meteorological stations – Hobart, Launceston, and Low Head Lighthouse – that are situated in the warmer and drier Eastern Tasmanian climatic zone (see also Cook et al. 1991, as communicated by John L. Daly of Tasmania).

familiar instrumental thermometer records that yield no information on centennial-scale climate variability. There are only three unambiguous findings favouring the 20th century as the warmest of the last 1000 years – the records from the Dyer Plateau, Antarctica, the Himalayas and Mongolia (Thompson et al. 1994; Thompson et al. 2000; D'Arrigo et al 2001). An important feature of Figure 3 is the large number of uncertain answers compared to the two prior questions. This feature is not easily explained, but it could be related to inaccurate calibration between proxy and instrumental data. Another interesting feature of the result is that the warmest or most extreme climatic anomalies in the proxy indicators often occurred in the early-to-mid 20th century, rather than throughout the century.

4.1. Glaciers – Worldwide

Broadly, glaciers retreated all over the world during the Medieval Warm Period, with a notable but minor re-advance between 1050 and 1150 A.D. (Grove and Switsur 1994). Large portions of the world's glaciers, both in the Northern and Southern Hemispheres, advanced during the 1300–1900 A.D. period (Grove 2001b; see also Winkler 2000). The world's small glaciers and tropical glaciers have simultaneously

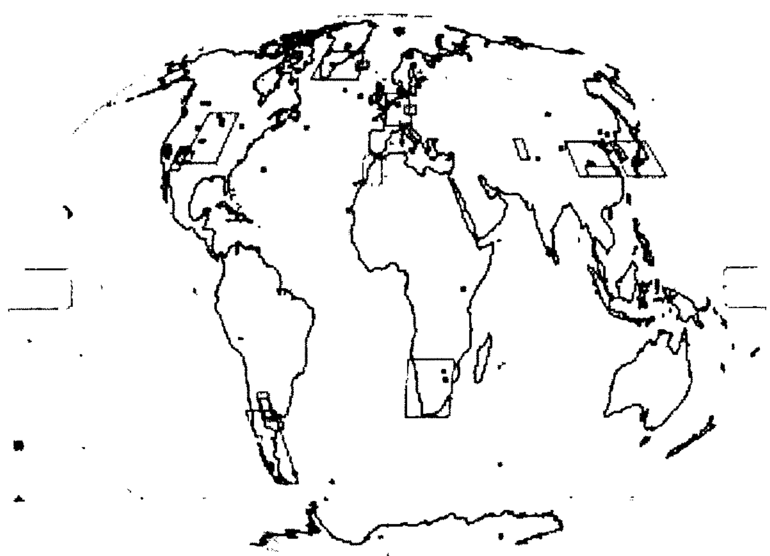


Figure 2. Geographical distribution of local answers to the following question: Is there an objectively discernible climatic anomaly during the Medieval Warm Period (800–1300 A.D.) in this proxy record? 'Yes' is indicated by red filled-squares or unfilled boxes, 'No' is indicated by green filled-circles and 'Yes? or No?' (undecided) is shown with blue filled-triangles or unfilled boxes.

retreated since the 19th century, but some glaciers have advanced (Kaser 1999; Dyurgerov and Meier 2000; D. Evans 2000). Kaser (1999) has reemphasized the key role played by atmospheric humidity in controlling the net accumulation and ablation of glaciers by modulating the sublimation and long-wave radiative forcing-feedback budgets in both dry and humid areas. So far, the proposition of 20th century warming being a natural recovery from the Little Ice Age, together with a possible amplification by anthropogenic CO_2, is plausible but not definitive (e.g., Bradley and Jones 1993; Kreutz et al. 1997; Kaser 1999; Beltrami et al. 2000; Dyurgerov and Meier 2000). By contrast, D. Evans (2000) discussed the possibility of recent widespread recession of glaciers as a glacioclimatic response to the termination of the Little Ice Age and commented that significant warming phases, especially those accompanied by relatively warm winters and cool summers, during interglacials may lead to the onset of another global glaciation.

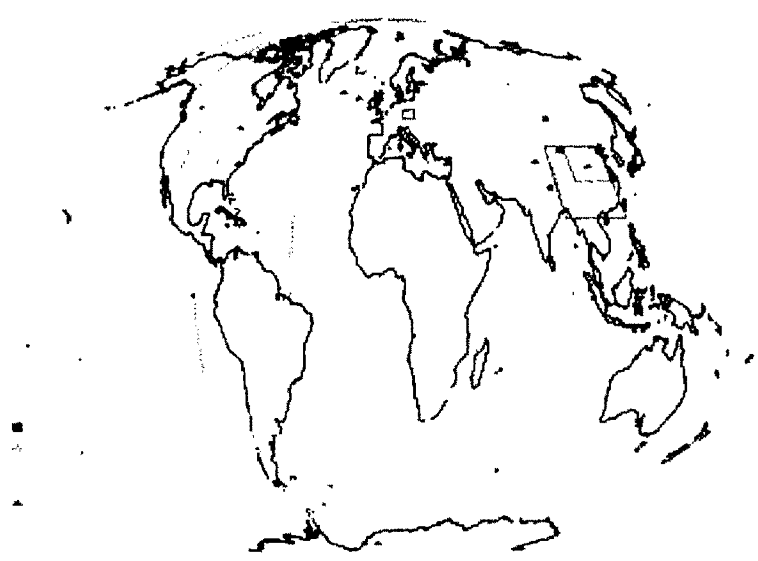


Figure 3. Geographical distribution of local answers to the following question: Is there an objectively discernible climatic anomaly within the 20th century that may validly be considered the most extreme (the warmest, if such information is available) period in the record? 'Yes' is indicated by red filled-squares, 'No' is indicated by green filled-circles or unfilled boxes and 'Yes? or No?' (undecided) is shown with blue filled-triangles or unfilled boxes. An answer of 'Yes*' is indicated by yellow filled-diamonds or unfilled boxes to mark an early to middle 20th century warming rather than the post-1970s warming.

Additional proxy records used here reveal that the climatic anomaly patterns known as the Medieval Warm Period (circa 800–1300 A.D.) and the Little Ice Age (1300–1900 A.D.) occurred across the world. The next two subsections describe detailed local changes over both the Northern and Southern Hemispheres.

4.2. Northern Hemisphere

A composite reconstruction of summer temperature (assuming a simple, uniform weighting of proxy records; see additional discussion in section 5.3) by Bradley and Jones (1993) showed that the 1530–1730 interval was the coldest period for the whole Northern Hemisphere and that the 19th century was the second coldest interval in the last 500 years.

4.2.1. Western Europe

Cold winters and wet summers prevailed during the Little Ice Age in Switzerland, where the most detailed and reliable information is available (Pfister 1995). A careful comparison of the Swiss and Central England (from Manley 1974) temperature records from 1659–1960 A.D. shows a general correspondence of climatic conditions between the two regions. In the Andalusia region of Southern Spain, rainfall appears to have alternated between wet and dry century-long spells (wet periods 1590–1649 and 1776–1937 A.D.; dry periods 1501–1589 and 1650–1775 A.D.) throughout the Little Ice Age, with no significant difference from the modern dry period of 1938–1997 (Rodrigo et al. 2000). Enhanced fluvial activity was documented in river basins of north, west and central Europe between 1250 and 1550 A.D. and again between 1750 and 1900 A.D. (A. T. Grove 2001). Over western Europe, Pfister et al. (1998) concluded that severe winters were less frequent and less extreme during 900–1300 A.D. than during 1300–1900 A.D. The mild-winter condition was hypothesized by Pfister et al. (1998) to have caused the northward migration of Mediterranean subtropical plants described by St. Albertus Magnus, who noted the abundance of pomegranates and fig trees in the 13th century around Cologne and parts of the Rhine valley. Olive trees, which, like fig trees, are also sensitive to prolonged periods of freezing, must have grown in Italy (Po valley), France and Germany, because a chronicler documented the damage to olive trees from the bitter frost in January 1234 A.D. Also, Lamb (1965) noted generally wet winters but drier summers for the lowlands in England, Ireland, Netherlands, Denmark, Sweden and northwest Germany from about 1200 to 1400 A.D. Those conditions are supported by documentary records that describe frequent flooding and storms around those regions during this transitional period between the Medieval Warm Period and Little Ice Age.

Was the warmth of the 1980s in western Europe exceptional or unusual? Lamb (1997a, page 386) observed that "even the great warmth of the years 1989/1991, hailed in some quarters as proof of the reality of the predicted global warming due to the enhancement of the greenhouse effect by increasing carbon dioxide and other effluents, requires the usual adjustments [i.e., from the expectation based solely on global warming model predictions]. ... it may also have a surprising analogy in the past to the remarkable warmth – well attested in Europe – of the year 1540, shortly before the sharpest onset of the so-called Little Ice Age. Pfister records that for several

decades before 1564 the climate in Switzerland – and this seems to be in line with the implications of other European chronicles – was on average about 0.4°C warmer, and slightly drier than today. The summers in the 1530s were at least as warm as in the warmest ten years of the present century, between 1943 and 1952. And the year 1540 outdid the warm dry year 1947 appreciably. From February till mid-December rain fell in Basle on only ten days. And young people were still bathing in the Rhine on the Swiss-German border at Schaffhausen in the first week of January 1541 after a ten-months-long bathing season. The warm anomaly of 1540 is the more remarkable because the weather then became severely wintry, and spring came late in 1541. Moreover, only twenty-four years later the 1564–5 winter was one of the longest and severest in the whole millennium in most parts of Europe and marked the arrival of the most notable cold climate period of the Little Ice Age, with ten to twenty historic winters, very late springs and cool summers and advancing glaciers."

Updated weather reconstruction results for the Low Countries (the present-day Benelux region) suggest that a meaningful answer to the question of whether the 20th century has the warmest extremes may be quite elusive until the seasonal dependence and resolution of a proxy-climate relation can be affirmed. For example, van Engelen et al. (2001) demonstrated that when the historical reconstructed proxy series from about 800–2000 A.D. was calibrated to instrumental temperature records at De Bilt, 20th century winter temperatures may have been slightly higher[7] than the high winter temperatures of 1000–1100 A.D.; but the 20th century summer temperatures are neither unusual nor extra-ordinarily warm when compared to natural summer temperature variabilities during other times of the second millennium (see Figures 1 and 2 of van Engelen et al. 2001).

4.2.2. North Atlantic and other oceans

During the Little Ice Age, extensive areas around Mediterranean Europe and the North Atlantic, including western and northern Europe, Greenland and Iceland, experienced unusually cold and wet conditions, as well as many extreme weather events, including deluges, landslides and avalanches (Grove 1996; Ogilvie et al. 2000; A. T. Grove 2001). From various proxies, the climate over Iceland was mild from 870 to 1170 A.D., with cold periods setting in after 1200 A.D. Instead of being a period of unrelenting cold, however, Ogilvie (1984) emphasized that the most notable aspect of climate over Iceland during the 17th to 19th centuries, with its very cold decades during the 1690s, 1780s, 1810s and 1830s, was its large year-to-year variability (see also Ogilvie and Jónsson 2001). The colonization of Greenland's coastal area by the Vikings starting in 986 A.D. is well documented; and the generally mild and benign climatic conditions from about 800–1200 A.D. that helped to sustain the settlement are also well supported by ice core and borehole proxy information (Dansgaard et al. 1975; Dahl-Jensen et al. 1998). The Norsemen's 'Western Settlement' (around the Godthab district) was mysteriously abandoned sometime between 1341 and 1362

7 About 0.5°C – but such warming was also clearly initiated earlier in the 19th century and the quantitative information of the 20th century warmth is certainly within the margin of uncertainty of this careful reconstruction effort.

A.D., while the 'Eastern Settlement' (actually near the southernmost tip of west Greenland, around the Narssaq and Julianehab districts) died out between 1450 and 1500 A.D. (Grove 1996; Ogilvie et al. 2000).[8] The timing of the abandonment of the settlements coincided with a general cooling over Greenland, as established by both ice-core isotopic and borehole thermometry (Dansgaard et al. 1975; Stuiver et al. 1995; Dahl-Jensen et al. 1998). From sediment cores near Nansen Fjord, East Greenland, Jennings and Weiner (1996) confirmed an initial cooling between 1270 and 1370 A.D., together with the most severe and variable climatic conditions around the East Greenland region from 1630–1900 A.D. The results of Ogilvie et al. (2000) and Ogilvie and Jónsson (2001) suggest that the overall climatic conditions in the North Atlantic (50–80°N; 0–60°W), especially near Iceland during the 20th century, including the 1970s–1990s, were neither unusual nor extreme.

In the Mediterranean basin, the island of Crete experienced many severe winters and prolonged droughts during the winter and spring seasons between 1548 and 1648 A.D. (Grove and Conterio 1995). In Morocco, the climate during the 16th, 17th and 18th centuries was generally more variable, with frequently drier conditions, than in the early to mid-20th century (Till and Guiot 1990). But no distinctive precipitation anomaly was observed for Morocco during the Medieval Warm Period, although just like conditions during the Little Ice Age, an episode of notable drought occurred from 1186–1234 A.D. Thus, precipitation anomalies for the Little Ice Age and Medieval Warm Period are not distinct from each other. For this reason, we designated the occurrence of a distinct climatic anomaly associated with the Medieval Warm Period for Morocco in Figure 2 as uncertain or 'Yes?'.

Distinctly cooler conditions prevailed over the oceans – in the Carribean Sea by about 2–3°C (Winter et al. 2000) and in the Sargasso Sea by about 1°C (Keigwin 1996), especially during the 17th and 18th centuries relative to the present. Likewise, deMenocal et al. (2000) found that the subtropical North Atlantic's sea surface temperature off Cap Blanc of Mauritania also was cooler by 3–4°C between 1300–1850 A.D. than at present. Similarly, during the Medieval Warm Period, the Sargasso sea surface temperature was about 1°C warmer than it is currently, while the sea surface temperature off the coast of Mauritania (west Africa) was only marginally warmer than at present. Based on Mg/Ca paleothermometry of the calcitic shells of microfossils, Cronin et al. (2003) found large (about 2 to 4°C warming or cooling) and rapid (< 100 years) shifts of spring water temperature of the Chesapeake Bay at 2100, 1600, 950 (1050 A.D.), 650 (1350 A.D.), 400 (1600 A.D.) and 150 (1850 A.D.) years before present. This Chesapeake Bay proxy temperature record of Cronin et al. (2003) recognizes five distinct climatic periods: the 20th century warm period,[9] the early

8 Barlow et al. (1997) emphasized that both cultural and political factors combined to make the Norse Greenlanders at the 'Western Settlement' more vulnerable to the harsh climatic conditions.

9 Although Cronin et al. (2003) suggested that "recent decadal climate variability in the North Atlantic [with similar suggestion for their Chesapeake Bay record] region is extreme relative to long-term patterns may be in part anthropogenic in origin", we noted that these authors admitted to the possible bias from greater sampling resolution in the last two centuries and "rapid regional warming ≈1800 AD [at Chesapeake Bay] was accompanied by accelerated sea level rise at the end of the Little Ice Age, about 1750–1850 AD, thus preceding large-scale fossil fuel burning". Thus we assigned 'Yes?' as answer to Question 3 in Table 1.

Medieval Warm Period (from 450–900 A.D.), the conventional Medieval Warm Period (from 1000–1300 A.D), the early Little Ice Age (from 1450–1530 A.D.) and the late Little Ice Age (1720–1850 A.D.) High-resolution coral skeletal $\delta^{18}O$ and Sr/Ca ratio records from Bermuda indicated sea-surface temperature standard deviations of about ±0.5°C on interannual and ±0.3°C on decadal timescales during the 16th century, the ranges of variability are comparable to estimates from modern 20th century instrumental data (Kuhnert et al. 2002). However, these sub-annual resolution coral proxy data also show that although there may be large-scale climate signals like the North Atlantic Oscillation detectable at Bermuda, no correlation can be found with other Northern-Hemispheric-wide proxy reconstructions (i.e., in Kuhnert et al.'s case, they compared with Mann et al. 1998's temperature series) because of large spatial differences in climate variability. From sedimentary concentrations of titanium and iron, Haug et al. (2001) inferred a very dry climate for the Cariaco Basin during the Little Ice Age and relatively wetter conditions during the Medieval Warm Period.

Over the equatorial Central Pacific, around the NINO3.4 (5N–5°S; 160°E–150°W) region, Evans et al. (2000), in their skillful reconstruction of the ENSO-like decadal variability of the NINO3.4 sea surface temperature (SST), found an apparent sustained cool phase of the proxy NINO3.4 SST variability from about 1550 A.D. to approximately 1895 A.D., thereby extending the geographical area covered by the Little Ice Age Climate Anomaly. Evans et al. (2000) also added that the reconstructed NINO3.4 decadal-scale SST variability prior to the 17th century is similar to that of the 20th century, thus suggesting that recent 20th century Pacific ocean decadal changes are neither unusual nor unprecedented.

From the analysis of $\delta^{18}O$ (for proxy of sea surface salinity) and Mg/Ca (for proxy of sea surface temperature) compositions of planktonic foraminefera *Globigerinoides ruber*[10] from cores retrieved from deep ocean near the eastern edge of the Indonesian archipelago, Stott (2002) and Stott et al. (2002) confirms that sea surface temperature and salinity around this area of Western Pacific Warm Pool were significantly anomalous around the Medieval Warm Period and the Little Ice Age. During the peak of the Medieval Warm Period from about 900 to 1100 AD, the sea surface temperatures were estimated to reach as high as 30°C, although Stott (2002) emphasizes that the primary signal of the climatic fingerprint in this deep sea core is manifested through the sea surface salinity rather than sea surface temperature. Stott (2002) further suggests that the warm and more saline sea surface condition during the Medieval Warm Period and the cooler and less saline sea surface condition during the Little Ice Age are not unique throughout the Holocene but instead represent a pattern of millennial climate variability in the Western Pacific Warm Pool region.

4.2.3. Asia and Eastern Europe

From 49 radiocarbon-dated subfossil wood samples, Hiller et al. (2001) determined that

10 This surface-dwelling species is abundant under warm summer surface waters of the Western tropical Pacific while the species *Globigerinoides sacculifer* is noted by Stott et al. (2002) to be present in smaller amount because it cannot form shells at temperature higher than 27°C – thus was analyzed to deduce sea surface temperatures during the cooler winter months for the Warm Pool region.

the alpine tree-limit on the Khibiny low mountains of the Kola Peninsula was located at least 100–140 meters above the current tree-limit elevation during the relatively warmer time between 1000 A.D. and 1300 A.D. The summer temperatures corresponding to the tree-line shift during this warm time are estimated to have been at least 0.8°C warmer than today. Based mostly on documentary evidence, Borisenkov (1995) noted that Little Ice Age conditions began as early as the 13th century in Russia, with the characteristic of frequent climate extremes both in terms of severe winters, rainy and cool summers, and sustained droughts (up to a decade long). Middle Russia (around 50–60°N and 30–50°E) seems to have experienced its coolest winters around 1620–1680 A.D., its coolest summers and springs around 1860–1900 A.D., and distinctively warm conditions during the first half of the 16th century, similar to conditions for western Europe described above. In addition, ground surface temperature histories deduced from boreholes around the Czech republic suggest that winters during 1600–1700 A.D. were the most severe since at least 1100 A.D. (Bodri and Cermák 1999). The temperature-depth borehole records also yield a clear signature of an anomalously warm period for central Bohemia, especially around 1100–1300 A.D.

Yadav and Singh (2002) note that climate over the Himalaya region tends to vary greatly over short distances because of forcing by topography. However, temperature variations may have a better spatial coherence when compared to precipitation changes over these high-elevation areas. Based on a network of twelve tree-ring width chronologies from the western Himalaya region of India, Yadav and Singh (2002) found that the warmest spring temperatures of the 20th century around 1945–1974 A.D. were not the warmest of the last four centuries in their tree-ring proxy temperature record, nor was the character of 20th century warming found to be unusual. Esper et al. (2002b), in a study based on more than 200,000 ring-width measurements from 384 trees over 20 individual sites extending over the Northwest Karakorum of Pakistan and the Southern Tien Shan of Kirghizia, compiled a 1300-year long record of proxy temperature data that resolved decadal-scale variations. This record shows that the warmest decades since 618 A.D. in Western Central Asia occurred between 800 and 1000 A.D., while the coldest periods occurred between 1500 and 1700 A.D.

In China, Bradley and Jones (1993) found that the mid-17th century was the coldest period of the last millennium. New China-wide composite temperature averages recompiled by Yang et al. (2002) confirm this fact about the coolest period during the Little Ice Age in China. Yang et al.'s (2002) records further suggest the warm period in China from 800–1400 A.D., which roughly corresponds to the Medieval Warm Period defined in this paper. Yang et al. (2002) also note that large regional temperature variations are found for the warm period in China – in Eastern China and in the northeastern Tibetan Plateau, the warm conditions prevailed from 800–1100 A.D., while in Southern Tibetan Plateau, the warmest period occurred in 1150–1380 A.D. In contrast, Yang et al. (2002) found that the cool conditions during the Little Ice Age are more homogeneous and consistent among various climate proxies. Although not an emphasis in their work, Yang et al. (2002) further noted, from their "weighted" reconstruction curve, that the warmest period in China of the last two millennia occurred in 100–240 A.D. surpassing even the warming of the 20th century.

In northeastern China, frequent occurrences of extremely dry conditions prevailed during the 16th and 17th centuries (Song 2000). The dry conditions returned again in the 20th century and now cover a wider area (with indications of an increasing number of days with no discharge from the Yellow River; but these 20th century events are likely to be confused with other man-made factors). Based on a combination of subdecadal (< 10 years) and longer-term (> 50 years) isotopic records from stalagmites retrieved from Buddha Cave (80 km south of Xian, central China), Paulsen et al. (2003) showed that although there were general cool and warm period during the Little Ice Age interval and Medieval Warm Period, respectively, the patterns of precipitation around the area were considerably more variable. For example, the region around Buddha Cave were generally dry from ca. 1640–1825 AD but the interval 1475–1640 AD were a relatively wetter period. Based on a homogeneous set of typhoon records from 1470–1931 A.D., Chan and Shi (2000) documented the notably larger number of land-falling typhoons over Guandong Province in the early-to-mid 19th century. Using a $\delta^{18}O$ proxy record from peat cellulose with 20-year resolution and various Chinese historical records, Hong et al. (2000) showed the general cooling trend in surface air temperature during the Little Ice Age interval in northeastern China. Hong et al. found three of the coolest minima in the record centred around 1550, 1650 and 1750 A.D. An obvious warm period peaked around 1100 to 1200 A.D., coinciding with the Medieval Warm Period. The study of documented cultivation of *Citrus reticulata Blanco* (a citrus tree) and *Boehmeria nivea* (a perennial herb), both subtropical and temperature-sensitive plants, during the last 1300 years showed that northern boundaries for these plants had shifted and expanded; their northernmost location was reached around 1264 A.D. (Zhang 1994). Zhang then deduced that temperature conditions in the 13th century around central China must have been about 1°C warmer than at present. Ren (1998) found further evidence from a fossil pollen record at Maili Bog, northeast China, that summer monsoon rainfall from 950–1270 A.D. must have been generally more vigorous in order to explain the high deposition of several pollen taxa, which are (otherwise) unexplainable by human activity at those times.

Based on less precise climate proxies like cherry-blossom-viewing dates, lake freezing dates and historical documentation of climate hazards and unusual weather, Tagami (1993, 1996) found that a warm period prevailed between the 10th and 14th centuries, and a cold period between the late 15th and 19th centuries, over large parts of southern Japan. From a study of the number of days with snowfall relative to days with rainfall, Tagami (1996) concluded that the 11th and 12th centuries were unusually warm. During the Little Ice Age, Japanese summers were relatively cool from the 1730s to 1750s, in the 1780s, from the 1830s to 1840s, and in the 1860s, while winters were cold from the 1680s to 1690s and in the 1730s and 1810s. From the tree-cellulose $\delta^{13}C$ record of a giant Japanese cedar (*Cryptomeria japonica*) grown on Yakushima Island of southern Japan, Kitagawa and Matsumoto (1995) inferred a temperature of 2°C below average from A.D. 1600 to 1700 and a warm period of about 1°C above average between A.D. 800 and 1200.

4.2.4. North America

Overall, the composite summer temperature anomaly of Bradley and Jones (1993)

shows that, over North America, mean temperature of the 15th–17th centuries was 1°C cooler than the average of the reference period 1860–1959.

Over the southern Sierra Nevada, California, Graumlich (1993) found that the coolest 50-year interval in her 1000-year tree-ring proxy record occurred around 1595–1644 A.D., while the wettest 50-year period was 1712–1761 A.D. Those occurrences are consistent with our definition of a discernible climatic anomaly associated with the Little Ice Age interval of 1300–1900 A.D. Ely et al. (1993) noted from river records in Arizona and Utah that the most extreme flooding events occurred during transitions from cool to warm climate conditions, especially during the late 1800s to early 1900s. For the Central U.S.A. (33–49°N and 91–109°W), drought episodes were noted for the 13th–16th centuries (from data compiled by Woodhouse and Overpeck 1998). These droughts were of longer duration and greater spatial extent than the 1930s–1950s drought (i.e., the 'Dust Bowl' drought). Additionally, both Yu and Ito (1999) and Dean and Schwalb (2000) identified cycles of aridity lasting about 400 years from lake records of the Northern Great Plains, where the last dry condition peaked around 1550–1700 A.D. Based on the tree ring proxy of hydroclimatic conditions in Southern Manitoba, George and Nielsen (2002) found that the Red River basin had experienced extremely dry conditions between 1670 and 1775 AD. These authors also concluded that the multidecades-scale change of the hydroclimate across the northeastern Great Plains during the last 600 years had been remarkably coherent upon comparing their tree-ring results with the limnological records from North Dakota and Minnesota.

From an extensive collection of multiproxy evidence, Stine (1998) concluded that during the Medieval Warm Period prolonged intervals of extreme drought affected California, the northwestern Great Basin, and the northern Rocky Mountains/Great Plains, while markedly wetter regimes persisted over the Upper Midwest/sub-arctic Canada and Southern Alaska/British Columbia regions. There was also a significant but brief interval around 1110–1140 A.D. when moisture conditions changed from dry to wet in California, the northwestern Great Basin, the northern Rocky Mountains/Great Plains, and from wet to dry in the Upper Midwest/sub-arctic Canada and Southern Alaska/British Columbia. The most likely explanation for this rapid and dramatic switch from wet to dry conditions around the Upper Midwestern U.S. around 1100 A.D. is the contraction and subsequent expansion of the circumpolar vortex. Summer polar fronts shifted significantly southward, stopping the penetration of moisture-laden air from the Gulf of Mexico (based on early ideas of Bryson et al. 1965). Stine (1998) added the requirement of a concomitant jet-stream change, from zonal to azonal, in order to explain the distinct observed differences of the moisture patterns between the Upper Midwest and Southern Alaska/British Columbia. Graumlich's (1993) reconstruction of summer temperature and winter precipitation from trees in the Sierra Nevada confirmed the overall warm and dry conditions for California during Medieval times, when two of the warmest and driest 50-year intervals occurred – at 1118–1167, 1245–1294 A.D. and 1250–1299, 1315–1364 A.D., respectively.

Hu et al. (2001), based on their high-resolution (multidecadal) geochemical analysis of sediments from Farewell Lake by the northwestern foothills of the Alaska

Range, also found pronounced signatures of the Medieval Warm Period around 850–1200 A.D. During the Little Ice Age, the surface water temperature of Farewell Lake fell to a low in 1700 A.D. that was estimated to be about 1.75°C cooler than at present. They also noted that colder periods were in general wetter (in contrast to drier conditions during the Little Ice Age in the Central United States region described above) than the warm periods in this part of Northwest Alaska. On the Yucatan Peninsula, prolonged drought episodes recur approximately every 200 years, with the two most significant recent peaks centred around 800 and 1020 A.D. (Hodell et al. 2001). Timings of these severe droughts also seem to fit several known discontinuities in the evolution of the Mayan culture.

4.3. Southern Hemisphere

Figures 1, 2 and 3 highlight the scarceness of Southern Hemisphere coverage by proxy climatic information through the second millennium A.D.

4.3.1. New Zealand

In New Zealand, the $\delta^{18}O$ concentration in a stalagmite record from a cave in northwest Nelson shows the coldest times during the Little Ice Age to be around 1600–1700 A.D., while exceptionally warm temperatures occurred around 1200–1400 A.D., in association with the general phenomenology of the Medieval Warm Period (Wilson et al. 1979). The cooling anomaly around 1600–1700 A.D., apparent in the $\delta^{18}O$ stalagmite record, coincides with the smallest growth rings (i.e., the coolest period) for silver pine (*Lagarostrobus colensoi*) from Mangawhero of North Island. However, at Ahaura, South Island, the smallest ring width index of the 600-year record occurred about 1500–1550 A.D. (D'Arrigo et al. 1998). Williams et al. (1999) issued important cautions concerning the interpretation of stable isotope data from New Zealand, especially the correctional functional relations among temperature, precipitation and $\delta^{18}O$ data (which are strongly influenced by oceans surrounding New Zealand) from Waitomo, North Island speleothems. The mean annual temperatures at Waitomo from 1430–1670 A.D. were deduced, based on the analysis of $\delta^{18}O$ data from Max's cave stalagmite, to be about 0.8°C cooler than at present.

4.3.2. South Africa

Tyson et al. (2000) showed, through isotopic measurements of a stalagmite, that the interior region of South Africa, near the Makapansgat Valley (eastern part of South Africa), had a maximum cooling of about 1°C around 1700 A.D compared to the present. This cooling corresponds well with the maximum cooling signal contained in a coral record from southwestern Madagascar (Tyson et al. 2000). Tyson and Lindesay (1992) demonstrated that the Little Ice Age in South Africa exhibited two major cooling phases, around 1300–1500 A.D. and 1675–1850 A.D., with a sudden warming interval occurring between 1500 and 1675 A.D. In addition, Tyson and Lindesay suggested a weakening of the tropical easterlies that increased the incidence of drought during the Little Ice Age in South Africa – with a relatively drier condition for the summer rainfall region in the northeast, but a wetter condition for the winter rainfall region near the coastal Mediterranean zone in the southwest. At Makapansgat Valley,

the Medieval Warm Period peaked with a temperature about 3–4°C warmer than at present around 1200–1300 A.D. (Tyson et al. 2000). The multiproxy review by Tyson and Lindesay (1992) showed evidence for a wetter South Africa after 1000 A.D., when forest and wetland become more extensive, including the development of a riverine forest in the northern Namib desert along the Hoanib river during the 11th–13th centuries.

4.3.3. South America

Over southern South America's Patagonia, the Little Ice Age's climatic anomalies, as deduced from tree ring records, were manifest as cold and moist summers with the most notable, persistent century-long wet intervals centred around 1340 and 1610 A.D. (Villalba 1994). From a multiproxy study of lacustrine sediments at Lake Aculeo (about 34°S; 50 km southeast of Santiago, Chile), Jenny et al. (2002) found a period of greatly increased flood events centred around 1400–1600 A.D. (and in three other intervals: 200–400, 500–700 and 1850–1998 A.D.), which could be interpreted as increased winter rains from enhanced mid-latitude westerlies that usher in more frontal system activities. In contrast, during the Medieval Warm Period, the southern Patagonia region at latitudes between 47–51°S became abnormally dry for several centuries before 1130 A.D. when water levels in several lakes (Lake Argentino, Lake Cardiel and Lake Ghio) around the area dropped significantly. Also, trees like the southern beech, *Nothofagus sp.*, grew as old as 100 years in the basin of these lakes before being killed by reflooding of the lakes (Stine 1994).

Slightly north toward the Central region of Argentina (around the Córdoba Province), Carignano (1999), Cioccale (1999) and Iriondo (1999) noted the prevailing conditions for the advancement of the Andean glaciers during the Little Ice Age, with two distinct cold and dry intervals around the 15th to 16th, and the 18th to the early-19th centuries. The significant climate aridification and deterioration in central Argentina (in contrast to the more humid conditions and increased flood frequency in central Chile near Lake Aculeo) during the Little Ice Age interval is supported by the formation of large, parabolic sand dunes 150–200 meters long, 60–80 meters wide, and 2–3 meters high in the Salinas Grandes basin (Carignano 1999). Meanwhile, the Mar Chiquita Lake was transformed into a swamp surrounded by dunes in the 18th century. Today, Mar Chiquita is the largest lake in Argentina, covering a surface area of 6000 km^2 to a depth of 13 meters (Iriondo 1999). The climatic conditions during the Medieval Warm Period around Central Argentina were generally warmer and more humid than at other times in the second millennium A.D., when the dune fields were conquered by lakes and the Mar Chiquita Lake expanded beyond its present dimensions. Precipitation exceeded current levels, and the mean local temperature may have been about 2.5°C warmer, perhaps because of the southward shift of the tropical climate belt into this area (Iriondo 1999). The northern part of Córdoba Province was invaded by the eastern boundary of the Chaco Forest, which is located hundreds of kilometres to the northwest today (Carignano 1999). Cioccale (1999) further noted evidence for human cultivation of hillside areas in Central Andes, Peru, at places as high as 4300 meters above sea level around 1000 A.D.

4.3.4. Antarctica

The last important source of geographical information for conditions during the Medieval Warm Period and the Little Ice Age in the Southern Hemisphere is obtained from glaciers, ice cores and sea sediments on and around Antarctica. Although many notable physical, biological and environmental changes have recently occurred there, especially around the Antarctic Peninsula during the last 50 years (e.g., Mercer 1978; Thomas et al. 1979; Rott et al. 1996; Vaughan and Doake 1996; Smith et al. 1999; Doran et al. 2002; Marshall et al. 2002), most of the 20th century changes contained in the proxy records discussed here cannot be considered extreme or unusual (see Figure 3, also Vaughan and Doake 1996; D. Evans 2000). For example, Vaughan and Doake (1996) deduced that a further warming of 10°C would be required to destabilize the Filchner-Ronne and Ross ice shelves that support the West Antarctic Ice Sheet. Under an even more extreme parametric study using a coupled thermo-mechanical ice sheet and climate model, Huybrechts and de Wolde (1999) concluded that their results do not support a catastrophic collapse or even strongly unstable behaviour of the marine-based ice sheet on West Antarctica. Extended studies based on geological proxy evidence appear to support the very-long-term stability of the Antarctic Ice Sheets (e.g., Kennett and Hodell 1995; Sugden 1996).

For the Little Ice Age, advances of glaciers on South Georgia Island (which is half-covered by glaciers[11]) began after the late 13th century, with a peak advancement around the 18th–20th centuries (Clapperton et al. 1989). Glacier retreats occurred after about 1000 A.D., which corresponds to the timing for the Medieval Warm Period. Baroni and Orombelli (1994) noted a similar scenario for glacier advances and retreats during the Little Ice Age and Medieval Warm Period for the Edmonson Point glacier at the Terra Nova Bay area of Victoria Land on the Antarctic continent (East Antarctica). The Edmonson Point glacier retreated in two distinct phases, around 920–1020 A.D. and 1270–1400 A.D., and then advanced at least 150 meters after the 15th century. Isotopic thermometry from ice cores at Dome C (74.65°S; 124.17°E; elevation 3240 meters) and Law Dome (66.73°S; 112.83°E; elevation 1390 meters) both indicate cooler and warmer anomalies for the Little Ice Age and Medieval Warm Period, respectively (Benoist et al. 1982; Morgan 1985). High-resolution records of magnetic susceptibility from deep sea cores (Domack and Mayewski 1999; Domack et al. 2001) drilled near the Palmer Deep site (64.86°S; 64.21°W) off the Antarctic Peninsula also show a marked increased in bio-productivity and corresponding decrease in magnetic susceptibility because of dilution of the magnetite, with a peak centred around 1000–1100 years BP. This observation probably implies warm temperatures and minimal sea-ice conditions, coinciding with the Medieval Warm Period. In the same record, Domack and colleagues found a decrease in bio-productivity and a corresponding increase in magnetic susceptibility owing to less dilution of the magnetic minerals by biogenic materials, from about 700 to 100 years BP. This time period corresponds to the Little Ice Age of the 14th to 19th century and is likely to have been accompanied by cool and windy conditions. A similar

11 Based on Clapperton et al. (1989) with updates from an October 8, 2002's email testimony of Mr. Gordon M. Liddle, Operations Manager, Government of South Georgia and the South Sandwich Islands.

interpretation of low magnetic susceptibility and high bio-productivity and, high magnetic susceptibility and low bio-productivity for the warmer and cooler climates during the Medieval Warm Period and Little Ice Age interval, respectively, by Khim et al. (2002) based on their analyses of the deep sea core A9–EB2 (455 cm long) recovered from the Eastern Bransfield oceanic basin (61.98°S; 55.95°W). Abundance analyses of Na^+ sea salt in the ice core from Siple Dome (81.65°S; 148.81°W) also confirms the Little Ice Age anomaly characterized by substantial variability in the strength of meridional circulation between 1400 and 1900 A.D. (Kreutz et al. 1997).

However, there also are indications for significant regional differences in climatic anomalies associated with the two phenomena in Antarctica. The temperature at Siple Station (75.92°S; 84.25°W; elevation 1054 meters) was relatively warm from the 15th to early 19th century (although there were also noticeable decade-long cooling dips centred around 1525 A.D., 1600 A.D. and 1750 A.D.; see Figure 29.7 of Mosley-Thompson 1995). The 400-year isotopic temperature inferred from a core at Dalinger Dome (64.22°S; 57.68°W; elevation 1640 meters) on James Ross Island off the Antarctic Peninsula also showed 1750–1850 A.D. to be the warmest interval, followed by a cooling of about 2°C since 1850 and continuing to 1980 (Aristarain et al. 1990). A recent borehole temperature reconstruction from Taylor Dome, East Antarctica (77.8°S, 158.72°E, elevation 2374 meters) reported the same inverted temperature anomalies, during which the Little Ice Age interval was about 2°C warmer, while the coldest temperature of the past four thousand years was reached around 1000 A.D. (Clow and Waddington 1999; note that we omitted these discussions from Table 1 and Figures 1–3 since their results are only preliminary and presented in a conference abstract).

Stenni et al. (2002) found that the Little Ice Age did not appear in their new Talos Dome ice core results as a long-lasting cold period. They noted a more persistent cold period from 1680 to 1820 A.D. but emphasized the centennial scale of warmer and cooler spells punctuating the Talos Dome record. In terms of the 20th century warming noted in other proxy records from the Southern Hemisphere, Stenni et al. (2002) remarked that the Talos Dome record showed no clear signature and that the warming in their record occurs before 1930 and another warming pulse seems to begin just after 1970. Stenni et al. also concluded that, when comparing their Talos Dome results with three other East Antarctica isotope records, there is little temporal synchronicity for the strongest Little Ice Age cooling phase around East Antarctica because of the influence of the local geography. But even at a single location, different proxies may give different climatic expressions as a result of differences in either the proxies' spatial and temporal sensitivities or genuine microclimatic phenomena (see further discussion in Stenni et al. 2002 for East Antartica). From all proxy records assembled here, we note there is no simple way to classify the contrasting warm/cold climatic anomalies in terms of a convenient division between East and West Antarctica, as has been hinted in some previous studies. Thus, the climatic anomaly patterns of the Little Ice Age and Medieval Warm Period around Antarctica remain to be fully revealed.

5. DISCUSSION

The widespread, but not truly global, geographical evidence assembled here argues for the reality of both the Little Ice Age and the Medieval Warm Period, and should serve

as a useful validation target for any reconstruction of global climate history over the last 1000 years. Our results suggest a different interpretation of multiproxy climate data than argued by Mann et al. (1998, 1999, 2000a). Since calibration of proxy indicators to instrumental data is still a matter of open investigation, it is premature to select a year or decade as the warmest or coldest of the past millennium. However, we now present a scientific examination on the quality of the Mann et al. (1998, 1999, 2000a) reconstruction, focusing on its limitations, especially because[12] these results are prominently featured and promoted in the Third Assessment Report of the Intergovernmental Panel on Climate Change (Albritton et al. 2001).

5.1. An examination of Mann et al.'s analyses and results

Mann, Bradley and Hughes (1998, 1999) and Mann et al. (2000a) conducted one of the most ambitious attempts to reconstruct global temperature variability and its pattern over the past millennium. Based on many long proxy records and their match with five leading spatial-temporal EOFs from modern surface thermometer records, Mann and colleagues developed a quantitative temperature history of the Northern Hemisphere dating as far back as 1000 A.D.

That non-local view and representation of climate variability is also echoed by Bradley et al. (2000). But the mathematics of EOFs introduces a potential and significant bias, as mentioned above. More importantly the non-local view of climate change has limited application to interannual variability. Yet, even for interannual variation, careful studies like Lau and Nath (2001) have shown that changes in heat anomalies at an open, maritime site in the Central Pacific are more likely responses to local variations in wind dissipation, while balances of energetics in a coastal region like the Gulf of Alaska are more dependent on non-local atmospheric advection of temperature and heat anomalies. As the interaction timescale increases, crucial but currently unresolved questions on the thermal and dynamical constraints of local geography and the nature of air-sea coupling will become more important (see Häkkinen 2000; Seager et al. 2000; Marshall et al. 2001 for updated discussions on the nature of air-sea-ice coupling for generating interannual, decadal and multidecadal climate variability over the North Atlantic). Questions on the validity of the locality paradigm will come into sharper focus as climatic changes on timescales on the order of about 50–100 years are pursued.

Several facts regarding Mann et al.'s reconstruction methodologies and their limitations are germane to these issues. First, Mann et al. (1998) stated the most fundamental assumption in their multi-proxy reconstruction effort is that all spatial patterns of climate variation over the last 1000 years precisely follow the observed global pattern of change in near-surface air temperature of the last 80–90 years or so.[13] Second, Mann et al. (1998) emphasized that their results find little skill in

12 This rationalization was suggested and recommended by an anonymous referee.

13 Mann (2002) subsequently cautioned that the "calibrated relation is determined from the 20th-century period, during which anthropogenic forcing played a prominent role. The approach could therefore yield a biased reconstruction of the past if the fundamental patterns of past temperature variation differ from those recorded in modern surface temperatures." (page 1481).

reconstructing the first eigenvectors prior to 1400 A.D., because no data exist for useful resolution of hemispheric-scale variability. Third, although 12 additional proxies were added to allow Mann et al. (1999) to reconstruct back to 1000 A.D., as opposed to 1400 A.D. in Mann et al. (1998), the positive calibration/variance scores are carried solely by the first principal component (PC #1), which consists of high-elevation tree growth proxy records from Western North America (Mann et al. 1999). This fact has led Mann et al. (1999) to report that the spatial variance explained by the distribution of their proxy "networks" in the calibration and verification process is only 5%, and that it is the time component, not the spatial detail, that is "most meaningful" for their millennial reconstruction results. (It is then easy to see that Mann et al's 1000-year reconstructed 'Northern-Hemispheric mean temperature' is dominated by relative changes in the western North America time series – compare Figures 2a and 2b in Mann 2001b). Mann et al. (1999) also specifically emphasized that their calibration/verification procedure fails if they remove the one crucial Western North American composite tree ring series from the list of 12 proxies.

In light of these limitations, the retrodiction of hemispheric-scale temperature changes from 1400 A.D. to 1000 A.D. is not robust, and no scientifically confident statements can be made about global temperature changes for the last 1000 years. Nevertheless, Mann et al.'s (1998, 1999, 2000a) key conclusions are that:

(a) the 20th century is 'nominally the warmest' of the past millennium, valid at least over the Northern Hemisphere.
(b) the decade of the 1990s was the warmest decade and 1998 the warmest year of the last millennium at 'moderately high levels of confidence.'[14]
(c) the notion of the Little Ice Age as a globally synchronous cold period can be dismissed.
(d) the notion of the Medieval Warm Epoch, according to Lamb (1965), applies mainly to western Europe and was not a global phenomenon (see our perspectives in section 6).

In contrast to the first claim above are the earlier borehole-heatflow temperature results of Huang et al. (1997), who utilized more than 6000 heatflow-depth measurements distributed worldwide to deduce a composite ground temperature record over the last 20,000 years. Over the last 1000 years, Huang et al. found that the composite ground temperature 500–1000 years ago was warmer by 0.1–0.5°C than the present.[15] After the early warmth, the temperature cooled to a minimum of 0.2–0.7°C

14 This claim was made by Mann et al. (1999) without comment on the likely association of 1998's global warmth to the well-noted 1997/98 El Nino event in that paper.

15 Our own private communications with SP. Huang and between M. MacCracken and H. N. Pollack, and in turn kindly shared with us by M. MacCracken on June 7, 2001, confirm that the warm feature during the Medieval Warm Period derived using this relatively "lower-quality" heatflow data, rather than direct borehole temperature profilings, is robust. In fact, H. N. Pollack carefully explained that "We do have a paper in Geophysical Research Letters (v. 24, n. 15, pp. 1947–1950, 1997) that uses lower-quality geothermal data (note carefully: this does not mean borehole temperature profiles!) from some 6000 sites, and this analysis does show a MWP when analyzed as a global dataset. When analyzed as separate high latitude (>45 degrees) and low latitude (<45 degrees), the amplitude is greater at high latitudes and smaller at low latitudes, but [the Medieval Warm Period feature is] still present."

below today's level about 200 years ago. Unlike tree-ring proxy climate results, borehole temperature reconstructions lose high frequency climate information, thus making the direct calibration and comparison with surface thermometer results difficult (see, however, an inter-calibration attempt by Harris and Chapman 2001).

Post-Mann et al. (1998, 1999) tree-ring reconstruction (re)analyses, like Briffa (2000) and Esper et al. (2002a), have also clearly shown evidence for a Medieval Warm Period that is at least as warm as the 20th century, for at least up to 1990. This is why many authors, including Broecker (2001), cautioned against any definitive conclusion on the nature of true climatic change from proxy records or from EOF mathematical reconstructions.

Three existing criticisms of, and significant challenges to, the conclusions of Mann and colleagues by several other researchers are:

(1) The majority of the tree-ring records used by Mann et al. (1999) have been standardized (see footnote 5). That process removes nonclimatic tree growth factors; and as a result, most of the climate variability information on timescales longer than about thirty years is lost (see Briffa and Osborn 1999; Briffa et al. 2001 and our deduction of the upper limit below). Briffa and Osborn also emphasized the significance of the lack of good time-resolution paleo-records in contrast to Mann et al.'s claim of a large number of independent datasets that can be used for multi-proxy reconstruction.[16]

(2) Strong evidence has been accumulating that tree growth has been disturbed in many Northern Hemisphere regions in recent decades (Graybill and Idso 1993; Jacoby and D'Arrigo 1995; Briffa et al. 1998; Feng 1999; Barber et al. 2000; Jacoby et al. 2000, Knapp et al. 2001) so that after 1960–1970 or so, the usual, strong positive correlation between the tree ring width or tree ring maximum latewood density indices and summer temperatures have weakened (referred to as "anomalous reduction in growth performance" by Esper et al. 2002a). The calibration period of Mann et al. (1998, 1999, 2000a) ended at 1980, while 20 more years of climate data post-1980 (compared to the 80 years length of their calibration interval, 1902–1980) exist. If the failure of inter-calibration of instrumental and tree growth records over last two to three decades suggests evidence for anthropogenic influences (i.e., from CO_2, nitrogen fertilization or land-use and land-cover changes or through changes in the length of growing seasons and changes in water and nutrient utilization efficiencies and so on), then no reliable quantitative inter-calibration can connect the past to the future (Idso 1989). Briffa and Osborn (1999) have also elaborated on the impact of unusual tree growth on the calibration procedure of tree-ring climate proxies (see additional discussions in Jacoby and D'Arrigo 1995; Briffa et al. 1998; Barber et al. 2000; Briffa 2000; Jacoby et al. 2000). This matter has largely been unresolved, which means that global or Northern Hemisphere-averaged

16 But see updated attempts by Briffa et al. (2001) and Esper et al. (2002a) to retrieve longer timescale climatic information from tree-ring data that are fundamentally limited by actual segment lengths of individual tree-ring series, dead or living, concatenated to produce the composite tree-ring chronology. In Esper et al.'s case, the segment lengths of their individual tree-ring series are about 200 to 400 years.

thermometer records of surface temperature cannot be simply attached to reconstructed temperature records of Mann et al, based mainly on tree-ring width, which cannot yet be reliably calibrated, to the latter half of the 20th century.

(3) Broecker (2001) tentatively concluded that the Medieval Warm Period was a global-scale event, although the hypothesized climate responses may be[17] anti-phased between the northern and southern high-latitude or polar regions. In terms of Broecker's hypothesis, the strength of the Atlantic ocean's thermohaline circulation oscillates naturally on a timescale of about 1500 years, based on the original findings of Bond et al. (1997, 1999) and its bipolar seasaw imprints on climate (see also the quote from Lamb [1982, 1997a] in section 2).

The 1902–1980 period of surface thermometer records adopted by Mann et al. (1999) as the calibration interval at best samples two to four *repeatable* or characteristic "multi-decadal" (say, 20-to-30-year) events. Therefore, the base spatial pattern adopted by Mann et al. (1998, 1999, 2000a) does not account for any relevant climate changes that may recur over 40-year and longer, e.g., centennial timescales. There may be skill in resolving > 40-year changes for limited regions like the North Atlantic for as far back as 1700 A.D. (e.g., as studied in Baliunas et al. 1997; Delworth and Mann 2000; Cullen et al. 2001), but a similar conclusion cannot be reached for global-scale changes spanning the last millennium. A direct comparison of > 40-year temperature variability by Esper et al. (2002a) confirmed that Mann et al's Northern Hemispheric mean reconstruction has significantly underestimated the multidecadal

17 The tentativeness arises because we do not know the precise mechanism of change Broecker proposed: it can be either thermohaline circulation mediated type of change or tropical ocean-induced changes with large amplification and northern-southern hemisphere synchronization effects through the water vapour feedback.

18 There is an internal inconsistency in the claim by Mann and Hughes (2002) that the differences between their results (Mann et al. 1998, 1999, 2000) and Esper et al. (2002a) may be partly explained by the fact that "[h]alf of the surface area of the NH temperature record estimated by Mann et al. lies at latitudes below 30°N, where as the Esper et al. estimate is based entirely on latitude above 30°N." Figure 9 of Mann et al. (2000) clearly show that the reconstruction record of Mann et al. yields similar amplitude and time variability for both the NH-wide and NH extratropical (30–70°N) averaged temperatures. Huang and Pollack (2002), by comparing their borehole temperature results with the synthetic subsurface temperatures generated by both Mann et al. and Esper et al. reconstructions, argued that it would require "an extraordinary contrast between the tropics and the extra-tropical continents at the hemispheric scale to account for the substantial difference between the negative transient predicted by the MBH [Mann et al.] reconstruction and the positive transients predicted by the ECS [Esper et al.] reconstruction and observed in boreholes." By further studying both the instrumental and reconstructed temperature dataset of Mann et al. (2000) [results of this independent checking are available from wsoon@cfa.harvard.edu], we find neither verification nor independent support for the claims by Mann (2002) and Mann et al. (2003) that "[n]early all of the proxy reconstructions are seen ... to be internally consistent (i.e., well within the uncertainties of the Mann et al. reconstruction). [sic.]" The argument by Mann and colleagues is that consensus to the wide variety of proxy results can be reached by "simply" re-scaling all proxy datasets according to their sensitivities to different (1) spatial sampling patterns, (2) seasons and (3) latitude bands. By contrast, we suggest that the large differences in the results from different proxy reconstruction efforts (i.e., Esper et al [2002] versus Mann et al. [1999] and Huang et al. [2000] versus Mann et al. [1999]) are real, owing to the differences in the long-term variance captured by the different proxies or internal biases of each proxy (with respect to their climatic information contents) or both.

and centennial scale changes[18] (see Figure 3 of Esper et al. 2002a). Another reason we do not accept the conclusions of the two most recent studies by Mann and colleagues in claiming that both GCM and proxy reconstructions have skill for the study of multidecadal climatic changes (see also contradictions of claims by Mann and colleagues in Collins et al. 2002) is the fact that various proxy-based reconstructions of the North Atlantic Oscillation (NAO) index back to 1675 demonstrated little verification skill, especially around the late 18th- to mid-19th-century as compared to NAO results based on available instrumental records (Schmutz et al. 2000). Evans et al. (2002) demonstrated similar problems in inter-calibration skill when comparing various paleo-reconstructions of ENSO-mode variability over the 19th century; and, hence, they cautioned about over-confidence in all proxy reconstructed versions of ground-truth. Zorita et al. (2003) examined the statistical reconstruction model adopted by Mann et al. (1998) in terms of important problems and questions related to (1) the available number of proxy indicators for estimating global-scale temperature, (2) how valid is it to apply the statistical reconstruction model that was calibrated at interannual time scale for estimating lower-frequency variability of past climate and (3) the influence of the limited spatial coverage of the current instrumental records on the calibration of the statistical reconstruction model.

Ogilvie and Jónsson (2001) have further noted that all current calibrations of proxies to large-scale instrumental measurements have been mainly valid over phases of rising temperature. The concern is that a different calibration response arises when the procedure is extended to an untested climate regime associated with a persistent cooling phase. Evans et al. (2002) worried about the reality of spurious frequency evolution that may contaminate a multiproxy reconstruction in which the type of proxy data changes over time and no sufficient overlap of proxy data exists for a proper inter-proxy calibration/validation procedure. In other words, each proxy may have its distinct frequency response function, which could confuse the interpretation of climate variability. Finally, another concern is the lack of understanding of the air-sea relationship at the multidecadal time scale, even in the reasonably well observed region of the North Atlantic (Häkkinen 2000; Seager et al. 2000; Marshall et al. 2001; Slonosky and Yiou 2001; JS. von Storch et al. 2001).

Taking all the physical criticisms and technical problems together, we conclude that the answers proposed for several key questions on climate behaviour of the past millennium in Mann et al. (1998, 1999, 2000a) are uncertain because of the unverifiable assumptions implicit in the mathematical extrapolation of the observed pattern of climatic changes – valid in the sampling of the 20–30 year scale of variability at most – to the full historical changes of the last 1000 years.

The Mann et al. (1998, 1999, 2000a) large-scale proxy temperature reconstruction is not capable of resolving the three specific questions we pose in this paper about the local reality of the Little Ice Age, Medieval Warm Period and 20th century warming (e.g., the mismatches between the local sea-surface temperature reconstruction from Bermuda and Mann et al. Northern-hemispheric proxy temperature shown in Figure 7 and additional comparative studies discussed on page 167 of Kuhnert et al. 2002). We have found that although the Mann et al. (1998, 1999, 2000a) reconstructed temperature seems to be well-calibrated for the annual-mean Northern-hemisphere-

wide (or globally averaged) instrumental temperatures, but we were not able to find any satisfactory calibrations for seasonal averages and/or for smaller regional averages (see also footnote 18). Thus, the composite time series of Mann et al. cannot yet be considered a realistic constraint for both timing and amplitude for global- or hemisphere-scale climatic changes of the past millennium, as further applied by, for example, Crowley (2000) to deduce *causes* of those changes in proxy-based reconstructions.

Briffa (2000, page 87) concluded that dendroclimatological records in general support "the notion that the last 100 years have been unusually warm, at least within the context of the last two millennia, [h]owever, this evidence should not be considered [un]equivocal [NB: our correction to Briffa's statement]." Later, Briffa et al. (2001), by adopting a new analysis procedure that seeks to preserve greater, long timescale variability (which shows a notable increase in variance at the 24–37 year time scale compared with a previous standardization procedure) in their tree ring density data than previously possible, stated that the 20th century is the globally warmest century of the last 600 years. This conclusion is consistent with the borehole reconstruction results of Huang et al. (2000). (Both Briffa et al. [2001] and Huang et al.'s [2000] new reconstruction did not extend back to 1000 A.D.) However, longer and more carefully-reconstructed tree-ring chronologies from Esper et al. (2002a) show that the Medieval Warm Period is indeed as warm as, or possibly even warmer than, the 20th century for at least a region covering the Northern Hemisphere extratropics from about 30°N to 70°N.

An important aspect of both the Briffa et al. (2001) and Esper et al. (2002a) studies is the new derivation of formal, time-dependent standard errors for their temperature reconstructions, amounting to about ±0.1 to 0.3°C from 1000 through 1960 (see also Jones et al. 1999; Jones et al. 2001). This assignment of standard errors contrasts with those assigned in Mann et al.'s (1999) annually-resolved series, where the uncertainties were assigned only for pre-instrumental data points in their original publication (their assumption of 'error-free' instrumental thermometer data is incorrect – see Jones et al. 1999, Folland et al. 2001 and the discussion of systematic adjustments and issues of surface temperature measurements in section 5.4). Over the full second millennium, Esper et al. (2002a) deduced a slightly larger range in their confidence limits after 1950 (compared to the pre-1950 interval extending back to 800 A.D.) and attributed those higher uncertainties to the anomalous modern ring-growth problem.

The accumulation of wide-spread proxies and the need to augment results like those from EOFs require a systematic re-examination of the qualitative results from many climate proxy indicators. The conservative view about standard errors (Briffa et al. 2001; Esper et al. 2002a) is adopted as a guide for a lower bound of errors (i.e., an error as large as 0.6–1.0°C for a confidence level of three standard deviations) in our analysis. This approach is guided by Jones et al. (2001), who emphasize the poor quality of paleoclimatic information over the Southern Hemisphere and say "it is dangerous to place too much reliance on these curves [NB: referring to the multiproxy summer temperatures for the Southern Hemisphere], because the associated errors are likely greater than those for the NH [Northern Hemisphere]." In addition, we ignore

all systematic errors.

5.2. On links between hemispheric-scale climate and tropical Pacific ENSO

The observed global warming of the 1990s as a decade, or even 1998 as a single year (which has been claimed to be the warmest of the Second Millennium), may be tied to strong El Nino-related events. Yet there is no clear sign that recent El Nino events are deterministically unusual compared to those of the last 350–400 years (Wunsch 1999; Cane 2001; Evans et al. 2002; or see Figure 14 in Mann et al. 2000a or Figures 17 and 18 in Mann et al. 2000b). To compare the physical extent of recent to past El Nino events, consider, for example, the historical results on the severity and wide-ranging impacts of the 1789–1793 ENSO by R. H. Grove (1998).[19] Kiladis and Diaz (1986) concluded, based on careful studies of historical and instrumental data, that the very-strong El Nino of 1982/83 was really not so exceptional in terms of its climatic anomalies when compared to the 1877/78 El Nino event.

The direct link between tropical ENSO events and global or hemispheric-scale climate parameters like temperature and rainfall is drawn from reasonably well-observed air and sea conditions and their hemisphere-wide teleconnection influences seen for the very-strong El Nino events of 1982/83 and 1997/98 (e.g., Harrison and Larkin 1998; Bell et al. 1999; Krishnamurti et al. 2000; Ueda and Matsumoto 2000; Hsu and Moura 2001; Kumar et al. 2001; Lau and Weng 2001). In that respect, it is noteworthy to contrast the *a posteriori* reasoning of Mann et al. (2000b) to achieve an internal consistency of their claim of the 1990s having been the warmest decade and 1998 the warmest single year of the past millennium (claimed by Mann et al. 1999) from the paleo-reconstructed Northern Hemispheric temperature with their Nino-3 (5°N–5°S; 90–150°W) sea surface temperature reconstruction, which showed no unusual ranking of those two events to mean that ENSO variability is very weakly coupled to global or hemispheric mean temperatures. That interpretation of Mann and colleagues is not internally consistent with what is known about robust teleconnection effects by ENSO during the 1980s and 1990s. A likely resolution might simply be that their paleo-reconstruction efforts are only calibrated by instrumental data up to 1980.

5.3. On criticisms of the Crowley and Lowery (2000) composite proxy curve

Although our approach and results are not directly comparable to those of Mann et al. (1998, 1999, 2000a), they can be compared with the results of Crowley and Lowery (2000). We have decided against a superposition of these diverse indicators of climate proxies because their individual sensitivities to temperature and other climatic

19 Historical evidence related to the 1789–1793 ENSO points to severe drought conditions around south India, Australia, Mexico, Southern Africa and regions around St. Helena (South Atlantic) and Montserrat (Carribean). But there is scarcely any indication of anything unusual in the 1791 global temperature anomaly pattern reconstructed by the multiproxy regression method of Mann et al. (2000b, Figure 25a) in those areas. On the other hand, the reconstruction of the rainfall anomaly field for the past 1000 years could be an entirely different problem altogether (Mann et al.'s effort focused solely on temperature reconstruction). For a careful discussion of recent observational evidence in terms of the 'trigger' and 'mature-enhancing' phases in the coupling between El Nino and the equatorial Indian Ocean, see for example Ueda and Matsumoto (2000).

variabilities are not well defined. Thus, the calibration steps of using a renormalization and an arbitrary arithmetic mean, and then calibration with instrumental data for *only selective time-intervals* (see below), as adopted by Crowley and Lowery (2000) to produce a composite curve, are simply biased correlation exercises. A selective set of proxy records, each with unequal spatial-temporal resolution and differing in climate sensitivity cannot be combined to produce a composite curve as readily as Crowley and Lowery (2000) assume.

The composite curve of Crowley and Lowery (2000) yields results inconsistent with its underlying proxies as well as those discussed here. The authors conclude that "[d]espite clear evidence for Medieval warmth greater than present in some individual records, the new hemispheric composite supports the principal conclusion of earlier hemispheric reconstructions and, furthermore, indicates that maximum Medieval warmth was restricted to two-three 20–30 year intervals [identified by the authors as 1010–1040, 1070–1105, and 1155–1190 A.D.], with composite values during these times being only comparable to the mid-20th century warm time interval."

Crowley and Lowery later proceed to recalibrate the composite, non-dimensional curve to hemispheric-mean thermometer temperatures by using selective intervals, namely 1856–1880 and 1925–1965 A.D. The interval of 1880–1920 was claimed to be contaminated by "anomalous tree-ring growth due to the 19th century rise in CO_2." In contrast, Bradley and Jones (1993) first produced their composite Northern Hemisphere temperature curve and then proceeded to improve its confidence by examining the potential problem of spatial sampling because "there are still extremely large areas for which we have no data." Bradley and Jones checked their composite results against the entire record of available instrumental summer temperature anomalies from about 1850 to 1980, as shown in Figure 7 of their paper. This approach is in sharp contrast to that of Crowley and Lowery, who calibrate their composite proxy curve based on limited areal coverage to Northern-Hemisphere-wide averaged instrumental data for only selective time-intervals and then claim the composite proxy temperature to be valid or relevant for the whole Northern Hemisphere and for the full time interval covered by the instrumental and proxy records.

The omission of the 1880–1920 period in the instrumental calibration is problematic and its explanation by Crowley and Lowery is insufficient. The anthropogenic CO_2 fertilization effect, suspected as an influence after the 1960s, could not already have been occurring between 1880 and 1920 and not afterward (see e.g., Knapp et al. 2001 for a discussion on the impact of the post-1960s CO_2 rise on western juniper growth rates under water-stressed conditions in Central Oregon[20]). More puzzling, Crowley and Lowery claim that only five (from the White Mountains of the Sierra Nevada, central Colorado, Ural Mountains of western Siberia, Qilian Shan of western China, as well as a 'phenological' record from East China) out of 15 of their proxy series were affected by this CO_2 fertilization effect, while four other tree ring proxy records (those from Jasper, Alberta; northern Sweden; the Alps of southern France; and the Black Forest in Germany) utilized in their composite curve were

20 The evidence for anthropogenic CO_2-fertilization seems to be much weaker in Tasmania and New Zealand (see e.g., D'Arrigo et al 1998).

unaffected (see, however, the notable examples of late-20th-century reduced tree growth in the forests of interior central Alaska and western Canada by Jacoby and D'Arrigo 1995; Barber et al. 2000)

A more promising explanation for a non-climatic growth response may be related to land-use, landscape and soil nutrient changes, rather than a direct and too early CO_2 aerial fertilization effect in the late 19- and early 20th century. Even more likely, the problem of the mismatch between instrumental data and Crowley and Lowery's composite curve around 1880–1920 could simply be a real difference between individual local proxy change and Northern-Hemisphere-averaged temperature variation that cannot be remedied by the *ad hoc* re-justification scheme proposed in Crowley and Lowery (2000).

The composite curve of Crowley and Lowery was calibrated against Northern Hemisphere-mean instrumental temperature anomalies and that calibration suggests a Northern Hemisphere temperature reconstruction. But equally well, the composite curve could be calibrated against global temperature to produce a similar claim of statistical association. After all, the correlation between the Northern Hemisphere-mean and the global-mean annual temperatures over 1866–2000 is about 0.94. However, such a process may shed no new information because modern thermometer data, when averaged over hemispheric scales, are relatively insensitive to regional details. Thus, the information from the largest-scale of change has, ultimately, very limited value for the practical problem of understanding local- and regional-scale changes. (Both Briffa et al 2001 and Esper et al. 2002a provide a similar discussion on the difficulty of distinguishing among large-scale spatial averages when calibrating regional proxy data).

As a result, the composite curve presented by Crowley and Lowery (2000) contains little physical information, especially for objective tests relative to the nature of the Little Ice Age, the Medieval Warm Period and the 20th century warm period.

5.4. On problems of calibrating proxy indices to instrumental data

Thus far, we have avoided discussion of attribution of suspected climatic factors to observed changes over the last 1000 years; there remain several barriers in the way of achieving this goal. Barnett et al. (1999) made an important point; namely, that it is *impossible* to use available instrumental records to provide estimates of multi-decadal and century-long natural climatic variations. Thus, paleo-proxies remain our only hope for assessing the amplitude and pattern of climatic and environmental change in the pre-human era. We argue with Barnett et al. (1999) that each proxy should be studied in terms of local change before several records can be combined for regional and larger spatial-scale analyses and interpretations. Our conclusion derives mainly from: (1) the real possibility of non-stationarity in the proxy-climate calibration to instrumental records, (2) the lack of adequate superposition rules, given the existence of variability in each type of proxy, and (3) the lack of a clear physical understanding of the multidecadal climate variability from theoretical or empirical studies.

Although instrumental temperature records are believed to have passed quality-control tests (Jones et al. 1999; Folland et al. 2001; Jones and Moberg 2003), most of the ship-based measurements of bulk sea surface temperature exhibit large systematic

adjustments. These adjustments range from 0.1°C to 0.45°C between 1856 and 1941 in hemispheric-scale averages, owing to the attempt to homogenize old water bucket-based measurements with modern ship engine intake measurements (see Figures 6 and 7 and the discussion on pp. 570–576 of Parker et al. 1995; Folland et al. 2001). Thus, difficulty seems unavoidable when merging measurements using instruments of different sensitivities and responses, and warrants a warning added to the phrase "calibration and independent cross-validation using instrumental data" in any reconstruction efforts (e.g., Mann et al. 1998, 1999, 2000a). The corrupted part of the early records is likely to set the ultimate constraint on the limited use of the verification procedure – for example, over the 1854–1901 interval in Mann et al. (1998, 1999, 2000a). Furthermore, Hurrell and Trenberth (1999) have shown significant differences among four sea surface temperature datasets,[21] even for climatologies as late as 1961–1990. Those differences not only have important consequences for the proxy-calibration process, but also for the interpretation of atmospheric circulation, moist convection, and precipitation fields over the tropics. In addition, Christy et al. (2001) found a significant relative warming of the decadal trend in bulk sea surface temperature compared to nighttime marine air (sampled on ship decks at a few meters height) and lower tropospheric temperature over the tropics (20°N to 20°S). Any calibration of paleo-proxy indicators to a "surface temperature" that does not distinguish between the air and the sea includes such differences as unquantifiable uncertainties.

Another significant problem is the indication that an anthropogenic influence may have already left its fingerprint on the recent growth of trees across the Northern Hemisphere. If this anthropogenic effect were present in tree ring data, then the calibration and verification procedure designed for extended paleoclimatic reconstructions would be significantly corrupted by further uncertainties (Idso 1989). Even with the convincing calibration of the proxy and instrumental data within their overlap interval, the calibration failure in the last 2 to 3 decades obscures climatic information from the proxy recorders. Another major concern is the inherently very long delay, including stases, between climatic forcing anomalies and responses. For example, in the biological and glaciological proxies a long temporal inertia exists in the forcing-feedback system (e.g., Bryson 1985; Cole 1985; Calkin et al. 2001; Hormes et al. 2001 for vegetational and glacial changes and their physical delays). The suggested time lags are as long as a few thousand years!

Enormous difficulties remain before an adequate sampling of historical changes can be amassed for conclusions on the largest spatial scales. Also, the assumption of global coherence is likely incorrect (and that assumption can be shown to be incorrect

21 Emery et al. (2001a, 2001b) elaborated on the distinction between bulk sea surface temperature measured at depths from 0.5 m to 5.0 m below sea surface from buoys and ships and "skin" sea surface temperature inferred from infrared satellite measurements that sample only the top 10μ•m of the oceanic layer. Overall, bulk minus skin sea surface temperature differences have mean values of about 0.3°C with an rms variability of up to 0.4°C, but the distinction is not a matter of constant adjustment to account for the cooler skin temperatures. The skin temperature has the important distinction of pertaining to the molecular layer that controls the air-sea exchanges of heat, momentum and gases.

even for the anthropogenic CO_2-influenced 20th century). This is why the core result of Mann et al., based mostly on the 600–1000 yr long tree-ring proxy records, through the EOF-calibrated pattern of temperature change over the 1902–1980 interval (or the later procedure using the 1902–1993 interval), does not address important questions about the context of the recent 20th century change relative to the variability of the last millennium. Refocusing on local changes from multiple proxy records can yield important information on the scope of 20th century changes relative to changes of past centuries.

Finally, a reiteration of the charge of Ogilvie and Jónsson (2001) is in order: "climate researchers should continue to seek to chart the climate of the past thousand years with a fresh approach rather than attempting to fit their findings into the convenient straightjacket of those hackneyed labels, the 'Medieval Warm Period' and the 'Little Ice Age'." In fact, Kreutz et al. (1997) and the follow-up effort by Mayewski and White (2002), based mainly on mismatches of the nature of 20th century climatic change between various proxy indicators (i.e., from their polar ice core analyses while citing sea surface records from the Sargasso Sea and Santa Barbara basin) and instrumental thermometers, suggest that *the Little Ice Age has not yet ended*. These authors argue it is possible that many components of the climate system, besides temperature, are still responding to perturbations from the Little Ice Age. Adopting this unique perspective, Mayewski and White remark that "When the recent rise in temperature seen in the Mann record [see our discussion in section 5.1 for details] is compared with our ice core-generated records of atmospheric circulation, a curious conclusion arises: Atmospheric circulation patterns appear to be within the range of variability of the LIA [Little Ice Age], but temperatures over the last few decades are markedly higher than anything during the LIA... We are forced to conclude that the LIA is not yet over and therefore human-induced controls on temperature are at play. While natural climate remains the baseline, human factors [the authors, Mayewski and White, are referring mainly to anthropogenic emissions of greenhouse gases and sulfur] may now be overpowering the trends that natural climate would follow if left undisturbed." But such an interpretation of the 20th century surface thermometer warming is similarly contentious. Karlén (2001), for example, notes that according to the Vostok ice core record of atmospheric carbon dioxide, the present concentration of atmospheric CO_2 is about 100 ppmv higher than it was during any previous interglacial during the last 400,000 years. Thus, if climate were to respond sensitively to carbon dioxide, global temperatures, or at least Vostok temperature, today ought to be considerably higher than previous interglacials. Yet evidence exists to suggest that the "present interglacial [at least for conditions around Vostok][22] has been about 2°C cooler than the previous one and the climate is now, in spite of the recent warming, cooler than it was at the beginning of this interglacial" (Karlén 2001).

22 Other paleoclimatic reconstructions, e.g., of tropical sea surface temperatures during the last and present interglacials, using the Mg/Ca ratios of foraminiferal shells from sediment near the Indo-Pacific Warm Pool region (around the Makassar strait, Indonesia), led to a similar conclusion about the relative warmth (about 1°C warmer) of the previous interglacial, ca. 120–124 kyr BP (Visser et al. 2003).

6. CONCLUSIONS

This paper presents a survey of site-specific paleoclimatic reconstructions, then considers whether they indicate that the Medieval Warm Period and the Little Ice Age were observed on broad area of the globe. We conclude that the Medieval Warm Period and Little Ice Age are widespread climatic anomalies, although we emphasize the complex nature of translating the proxy changes into convenient measures like temperature and precipitation as well as confirming their spatio-temporal representation and resolution.[23]

The procedures and emphases of our study contrast with those of Mann et al. (1998, 1999, 2000a), whose results are mainly a mathematical construct. Also in contrast to them, our assessment maintains the wide view of including many local climatic and environmental changes over the last 1000 years rather than relying on a mathematical filter to sieve only temperature changes. The wider views may be more appropriate when one seeks a broader perspective on the nature of climatic and environmental changes of the past millennium (e.g., Bryson 1985).

Mann et al. (1999) suggest that there has been a misleading speculation: "that temperatures were warmer [than current 20th century global warming] even further back, ≈1000 years ago – a period described by Lamb (1965) as the Medieval Warm Epoch (though Lamb, examining evidence mostly from western Europe *never* [emphasis added] suggested this was a global phenomenon)." A similar statement has also been made by Bradley (2000).

We correct those claims that misrepresent Lamb's statements – the restatement by Mann, Bradley and colleagues is contrary to Lamb's original statements and published ideas (in several of H. H. Lamb's popular and semi-technical books; e.g., Lamb 1977; Lamb 1982; Lamb 1997b). (A useful verification of this rhetorical confusion about whether Lamb utilizes any evidence from outside of western Europe may be found in Lamb [1963], a preamble to Lamb's [1965] more well-cited paper, where Lamb discussed evidence for climatic anomalies from all over the world. See also the first quote in the Introduction to this paper.) However, criticism on the actual quality of data at Lamb's time is quite another matter; see, for example, updated comments in Pfister et al. (1998).

Unfortunately, current knowledge of Earth's climate system does not yield quantitative and conclusive answers on many straightforward questions regarding the geographical nature and physical causes of surface temperature or precipitation changes over the last 1000 years. Note also that the adopted period of 1000 years is strictly a convenience that merits little scientific meaning.

---

23 In this sense, we are forced to reject the temptation to come up with a "best-guess" depiction of the large-scale or hemispheric-scale averages of the temperature anomalies. A practical reason is the potentially large sampling errors introduced by various kinds of inhomogeneity. One useful quantitative estimate by Sakamoto and Masuda (2002) showed that when there is no *a priori* choice to the selection of the spatial distribution of the proxy data points across the globe, the difference between a 100-point average *surface* temperature (with actual accounting of the elevation at each of the 100 irregularly-distributed points given the current topography and climatology) and a 100-point average *sea-level* temperature can be more than 2°C.

Climate proxy research does yield an aggregate and broad perspective on questions regarding the reality of the Little Ice Age, the Medieval Warm Period and the 20th century surface thermometer global warming. The picture emerges from many localities that both the Little Ice Age and Medieval Warm Period are widespread and perhaps not precisely timed or synchronous phenomena, easily within the margin of viewpoints conceived by Bryson et al. (1963), Lamb (1965) and numerous other researchers like J. Grove (1996, 2001a, 2001b). Our many local answers confirm that both the Medieval Climatic Anomaly and the Little Ice Age Climatic Anomaly are worthy of their respective labels. Furthermore, thermometer warming of the 20th century across the world seems neither unusual nor unprecedented within the more extended view of the last 1000 years. Overall, the 20th century does not contain the warmest or most extreme anomaly of the past millennium in most of the proxy records.

However, it is also clear that human activity has shaped almost every aspect of past environmental and climatic changes on local and regional spatial scales (perhaps on scales as small as 10 to 1000 km^2 for precipitation and 10^4 to 10^5 km^2 for temperature; see e.g., Messerli et al. 2000). For example, palynological analyses of two cores from the Huanghe (Yellow River) delta, with evidence for a major reduction in arboreal pollen [*Quercus* (*Lepidobalamus*)] followed by sudden increases in sediment discharge, conifer [*Pinus* (*Diploxylon*)] and buckwheat pollen [*Fagopyrum*] around 4 kyr to 1.3 kyr BP, suggested significant human-induced vegetational changes through deforestation and agricultural cultivation (Yi et al. 2003). Lawton et al. (2001) showed how the deforested areas of tropical lowlands can, in combination with favourable topographical conditions and altered atmospheric air flow across the landscape, significantly raise the bases of convective and orographic clouds around the Monteverde montane cloud forests of Costa Rica during the dry season, and thus drastically impact local ecosystems. However, see A. T. Grove (2001) for a clarification on the imprecise and misleading claim of the dominant role played by human activity (deforestation, agricultural expansion and population growth) on geomorphological changes (soil erosion or rapid sedimentation in river valleys and deltas to form the 'younger fill' of the Mediterranean Europe's fluvial terraces) in Mediterranean Europe, instead of the more powerful influences from Alpine glacier advances associated with the Little Ice Age.

Yet, subjective exercises to superpose the two not-entirely-compatible instrumental temperature and proxy climate time series need a lot more attention. It might seem surprising or frustrating that paleoclimatic reconstruction research has not yet provided confident and applicable answers to the role of anthropogenic forcing on climate change. This point is particularly sharp when considering the fact that even though some proxy records (e.g., those from Overpeck et al. 1997) show unprecedented 20th century warmth with most of the increase occurring in the early to mid-decades of the 20th century, when the amount of anthropogenic CO_2 in the air was less than 20–30% of the total amount there now. Unless there are serious flaws in the timing of the early-to-middle 20th century surface thermometer warming, or unknown anthropogenic mechanisms that caused a large amplification of surface temperature of the then-small increase in anthropogenic atmospheric CO_2, then the early part of the

20th century warming must be largely dissociated from anthropogenic CO_2 emissions. Other anthropogenic factors still need to be studied on a case by case basis.

Thus, a resolution of the pattern and amplitude of natural climate variability on multidecadal and centennial timescales through a multiproxy approach remains extremely important. The results may help quantify the relative apportionment of natural versus anthropogenic factors of recent climate change (i.e., the last three decades or so). The other avenue to quantify natural and anthropogenic climate variability using sophisticated general circulation models (GCMs) still suffers from the following problems: (i) the GCMs' underestimation of climate variance on multidecadal and centennial timescales (e.g., Barnett et al. 1999; von Storch et al. 2001; Collins et al. 2002), (ii) large differences among model-generated variability on both local and regional scales (e.g., Räisäanen 2001), and (iii) unrealized climate variability if based only on one or two realizations[24] of a forcing scenario (e.g., Delworth and Knutson 2000; Andronova and Schlesinger 2001) or because of the GCM's inability to account for certain biochemical and biophysical feedbacks and nondeterministic component of the earth's climate system (Idso 1998; Ou 2001; Pielke 2001). JS. von Storch et al. (2001), for example, present a critical discussion on the lack of confidence in the representation by current GCMs of low-frequency climate variability related primarily to the inability of models to resolve small-scale oceanic eddies. In addition, natural climate variations on multidecadal, centennial and millennial timescales could be highly non-stationary with complex spatio-temporal phasing (i.e., with part of such characteristics documented in the present study) that would be difficult for GCMs to emulate robustly. Such natural climatic factors likely operate in the real world together with further complications such as forcing by anthropogenic greenhouse gases and multi-component (rather then sulfate aerosol alone) aerosols. From the longer term perspective offered by paleoclimatic studies, one can at least conclude that large and rapid climatic and environmental changes have been common over the past millennium. Such natural changes impacted human society significantly, so further climate research, including adaptation strategies, rather than mitigation, seems to be pertinent (e.g., chapter 18 of Lamb 1982; Pielke 1998).

The quest in paleoclimate reconstruction efforts is to decipher and understand the physical mechanisms associated with the widest possible range of climate variabilities. That goal requires careful and systematic observation of the present-day Earth. The logistic and technical feasibility for the important objective and bias-free strategy to detect global warming due to the enhanced atmospheric concentration of carbon dioxide has been elaborated by Goody et al. (1998) and Keith and Anderson (2001).

ACKNOWLEDGMENTS

This work was supported by funds from the Air Force Office of Scientific Research (Grant AF49620-02-1-0194), the American Petroleum Institute (Grants 01-0000-4579 and 2002-100413) the National Aeronautics and Space Administration (Grant NAG5-

---

24 For example, Cherchi and Navarra (2003) showed that a minimum ensemble size of 16 members is needed in order to capture the observed interannual (1979–1993) variability for their focus on the nature of sea-surface-temperature-forced responses of the South Asian summer monsoon system.

7635), and the National Oceanic and Atmospheric Administration (Grant NA96GP0448). The views expressed herein are those of the authors and are independent of sponsoring agencies.

We have benefited greatly from the true and kind spirit of research communications (including constructive criticisms and a preview of their thoughts) with the late Jean Grove (who passed away on January 18, 2001), Dave Evans, Shaopeng Huang, Jim Kennett, Hong-Chun Li, Lowell Stott, Yoshio Tagami and Taiping Zhang. We thank John Daly and Diane Douglas-Dalziel for their unselfish contributions of many excellent collections of references. We are also very grateful to Maria McEachern, Melissa Hilbert, Barbara Palmer and Will Graves for invaluable library help and both Philip Gonzalez and Lisa Linarte for crucial all-around help.

We thank the three anonymous referees for constructive comments.

REFERENCES

Albritton DL and 58 others (2001) Summary for Policymakers. In Houghton JT et al. (eds) Climate Change 2001: The scientific basis, Cambridge University Press, p. 1–20

Ambaum MHP, Hoskins BJ, Stephenson DB (2001) Arctic oscillation or North Atlantic oscillation? J Clim 14: 3495–3507 [with corrigendum in J Clim 15: 553 (2002)]

Andrén E, Andrén T, Sohlenius G (2000) The Holocene history of the southwestern Baltic Sea as reflected in sediment core from the Bornholm Basin. Boreas 29: 233–250

Andronova NG, Schlesinger ME (2001) Objective estimation of the probability density function for climate sensitivity. J Geophys Res 106: 22605–22612

Aristarain AJ, Jouzel J, Lorius C (1990) A 400 years isotope record of the Antarctic Peninsula climate. Geophys Res Lett 17: 2369–2372

Arseneault D, Payette S (1997) Reconstruction of millennial forest dynamics from tree remains in a subarctic tree line peatland. Ecology 78: 1873–1883

Aykroyd RG, Lucy D, Pollard AM, Carter AHC, Robertson I (2001) Temporal variability in the strength of the proxy-climate correlations. Geophys Res Lett 28: 1559–1562

Baliunas S, Frick P, Sokoloff D, Soon W (1997) Time scales and trend in the Central England Temperature data (1659–1990): A wavelet analysis. Geophys Res Lett 24: 1351–1354

Barber VA, Juday GP, Finney BP (2000) Reduced growth of Alaskan white spruce in the twentieth century from temperature-induced drought stress. Nature 405: 668–673

Barclay DJ, Wiles GC, Calkin PE (1999) A 1119-year tree-ring-width chronology from western Prince William Sound, southern Alaska. Holocene 9: 79–84

Barlow LK and 8 others (1997) Interdisciplinary investigations of the end of the Norse Western Settlement in Greenland. Holocene 7: 489–499

Barnett T, Jones P (2000) Reply (to Bradley et al. 2000's comments). Bull Am Meteor Soc 81: 2990–2992

Barnett T and 10 others (1999) Detection and attribution of recent climate change: A status report. Bull Am Meteor Soc 80: 2631–2659

Baroni C, Orombelli G (1994) Holocene glacier variations in the Terra Nova Bay area (Victoria Land, Antarctica). Antarctic Sci 6: 497–505

Bell GD and 6 others (1999) Climate assessment for 1998. Bull Am Meteor Soc 80: S1 – S48

Beltrami H, Chapman DS, Archambault S, Bergeron Y (1995) Reconstruction of high resolution ground temperature histories combining dendrochronological and geothermal data. Earth Planet Sci Lett 136: 437–445

Beltrami H, Taylor AE (1995) Records of climatic change in the Canadian Arctic: Towards calibrating oxygen isotope data with geothermal data. Glob Planet Change 11: 127–138

Beltrami H, Wang J, Bras RL (2000) Energy balance at the Earth's surface: Heat flux history in Eastern Canada. Geophys Res Lett 27: 3385–3388

Benoist JP, Jouzel J, Lorius C, Merlivat L, Pourchet M (1982) Isotope climatic record over the last 2.5 ka from Dome C, Antarctica, ice cores. Ann Glaciology 3: 17–22

Bernabo JC (1981) Quantitative estimates of temperature changes over the last 2700 years in Michigan based on pollen data. Quat Res 15: 143–159

Betancourt JL, Grissino-Mayer HD, Salzer MW, Swetnam TW (2002) A test of "annual resolution" in stalagmites using tree rings. Quat Res 58: 197–199

Bianchi GG, McCave IN (1999) Holocene periodicity in North Atlantic climate and deep-ocean flow south of Iceland. Nature 397: 515–517

Biondi F, Perkins DL, Cayan DR, Hughes MK (1999) July temperature during the second millennium reconstructed from Idaho tree rings. Geophys Res Lett 26: 1445–1448

Black DE and 5 others (1999) Eight centuries of North Atlantic Ocean atmosphere variability. Science 286: 1709–1713

Blackford JJ, Chambers FM (1995) Proxy climate record for the last 1000 years from Irish blanket peat and a possible link to solar variability. Earth Planet Sci Lett 133: 145–150

Bodri L, Cermák V (1999) Climate change of the last millennium inferred from borehole temperatures: Regional patterns of climatic changes in the Czech Republic – Part III. Glob Planet Change 21: 225–235

Bond G and 9 others (1997) A pervasive millennial-scale cycle in North Atlantic Holocene and glacial climates. Science 278: 1257–1266

Bond G and 7 others (1999) The North Atlantic's 1–2 kyr climate rhythm: Relation to Heinrich events, Dansgaard/Oeschger cycles and the little ice age. In Clark PU, Webb RS, Keigwin LD (eds) Mechanisms of global climate change at millennial time scales, American Geophysical Union Press, vol 112, p. 35–58

Borisenkov YeP (1995) Documentary evidence from the U.S.S.R. In Bradley RS, Jones PD (eds) Climate Since A.D. 1500, Routledge, p 171–183

Bradley RS (2000) 1000 years of climate change Science 288: 1353, 1355

Bradley RS and 5 others (2001) The scope of Medieval Warming. Science 292: 2011–2012

Bradley RS, Hughes MK, Mann ME (2000) Comments on "Detection and attribution of recent climate change: A status report". Bull Am Meteor Soc 81: 2987–2990

Bradley RS, Jones PD (1993) 'Little Ice Age' summer temperature variations: Their nature and relevance to recent global warming trends. Holocene 3: 367–376

Briffa KR and 7 others (1992) Fennoscandian summers from AD 500: Temperature changes on short and long timescales. Clim Dyn 7: 111–119

Briffa KR and 7 others (1998) Trees tell of past climates: But are they speaking less clearly today? Phil Trans R Soc Lond B 353: 65–73

Briffa KR, Osborn TJ (1999) Seeing the wood from the trees. Science 284: 926–927

Briffa KR (2000) Annual climate variability in the Holocene: Interpreting the message of ancient trees. Quat Sci Rev 19: 87–105

Briffa KR and 6 others (2001) Low-frequency temperature variations from a northern tree-ring density network. J Geophys Res 106: 2929–2941

Broecker WS (2001) Was the medieval warm period global? Science 291: 1497–1499

Bryson RA (1985) On climatic analogs in paleoclimatic reconstruction. Quat Res 23: 275–286

Bryson RA, Bryson RU (1997) High resolution simulations of regional Holocene climate: North Africa and the Near East. In Dalfes HN, Kukla G, Weiss H (eds) Third Millennium BC Climate Change and the Old World Collapse, NATO ASI Series Vol. 149, Springer-Verlag, p. 565–593

Bryson RA and 39 others (1963) In Bryson RA, Julian PR (eds) Proceedings of the Conference on climate of the 11th and 16th centuries, National Center for Atmospheric Research Technical Notes 63-1, 103 pp.

Bryson RA, Irving WN, Larsen JA (1965) Radiocarbon and soil evidence of former forest in the Southern Canadian Tundra. Science 147: 46–48

Burns SJ and 5 others (2002) A 780-year annually resolved record of Indian Ocean monsoon precipitation from a speleothem from south Oman. J Geophys Res 107 (D20): 4434, doi:10.1029/2001JD001281

Calkin PE, Wiles GC, Barclay DJ (2001) Holocene coastal glaciation of Alaska. Quat Sci Rev 20: 449–461

Campbell ID, Campbell C, Apps MJ, Rutter NW, Bush ABG (1998) Late Holocene ≈1500 yr climatic periodicities and their implications. Geology 26: 471–473

Cane MA (2001) Book review – El Nino and the Southern Oscillation: Multiscale Variability and Global and Regional Impact, Diaz HF, Markgraf V (eds), Cambridge Univ Press. EOS 82: 556

Carignano CA (1999) Late Pleistocene to recent climatic chance in Córdoba Province, Argentina: Geomorphological evidence. Quat Int 57–58: 117–134

Chan JCL, Shi JE (2000) Frequency of typhoon landfall over Guangdong province of China during the period 1470–1931. Int J Clim 20: 183–190

Cherchi A, Navarra A (2003) Reproducibility and predictability of the Asian summer monsoon in the ECHAM4-GCM. Clim Dyn: in press

Christy JR and 5 others (2001) Differential trends in tropical sea surface and atmospheric temperatures since 1979. Geophys Res Lett 28: 183–186

Chu G and 6 others (2002) The 'Medieval Warm Period' drought recorded in Lake Huguangyan, tropical South China. Holocene 12: 511–516

Cioccale MA (1999) Climatic fluctuations in the Central region of Argentina in the last 1000 years. Quat Int 62: 35–47

Clapperton CM, Sugden DE, Birnie J, Wilson MJ (1989) Late-glacial and Holocene glacier fluctuations and environmental change on South Georgia, Southern Ocean. Quat Res 31: 210–228

Clow GD (1992) The extent of temporal smearing in surface-temperature histories derived from borehole temperature measurements. Palaeogeogr Palaeoclim Palaeoecol 98: 81–86

Clow GD, Waddington ED (1999) Quantification of natural climate variability in Central Greenland and East Antarctica using borehole paleothermometry (abstract). International Union of Geodesy and Geophysics XXII General Assembly Abstract Book: B249

Cohen AL, Tyson PD (1995) Sea-surface temperature fluctuations during the Holocene off the south coast of Africa: Implications for terrestrial climate and rainfall. Holocene 5: 304–312

Cole K (1985) Past rates of change, species richness, and a model of vegetational inertia in the Grand Canyon, Arizona. Amer Naturalist 125: 289–303

Collins M, Osborn TJ, Tett SFB, Briffa KR, Schweingruber FH (2002) A comparison of the variability of a climate model with paleotemperature estimates from a network of tree-ring indices. J Clim 15: 1497–1515

Cook ER and 7 others (1991) Climatic change in Tasmania inferred a 1089-year tree-ring chronology of Huon pine. Science 253: 1266–1268

Cook ER, Briffa KR, Meko DM, Graybill DA, Funkhouser G (1995) The 'segment length curse' in long tree-ring chronology development for palaeoclimatic studies. Holocene 5: 229–237

Cook ER, Buckley BM, D'Arrigo RD, Peterson MJ (2000) Warm-season temperatures since 1600 BC reconstructed from Tasmanian tree rings and their relationship to large-scale sea surface temperature anomalies. Clim Dyn 16· 79–91

Cowling SA, Skyes MT (1999) Physiological significance of low atmospheric CO_2 for plant-climate interactions. Quat Res 52: 237–242 (See additional research exchanges in Quat Res 53: 402–404 and Quat Res 53: 405–406)

Cronin TM, Dwyer GS, Kamiya T, Schwede S. Willard DA (2003) Medieval warm period, little ice age and the 20th century temperature variability from Chesapeake Bay. Glob Planet Change: in press

Crowley TJ (2000) Causes of climate change over the past 1000 years. Science 289: 270–277

Crowley TJ, Lowery TS (2000) How warm was the medieval warm period? Ambio 29: 51–54

Cullen HM, D'Arrigo RD, Cook ER, Mann ME (2001) Multiproxy reconstructions of the North Atlantic Oscillation. Paleocean 16: 27–39

Dahl-Jensen D and 6 others (1998) Past temperatures directly from the Greenland ice sheet. Science 282: 268–271

Dansgaard W and 5 others (1975) Climatic changes, Norsemen and modern man. Nature 255:

24–28

D'Arrigo RD and 6 others (1998) Tree-ring records from New Zealand: Long-term context for recent warming trend. Clim Dyn 14: 191–199

D'Arrigo RD and 8 others (2001) 1738 years of Mongolian temperature variability inferred from a tree-ring width chronology of Siberian Pine. Geophys Res Lett 28: 543–546

Davis OK (1994) The correlation of summer precipitation in the southwestern USA with isotopic records of solar activity during the Medieval Warm Period. Clim Change 26: 271–287

Dean WE, Schwalb A (2000) Holocene environmental and climatic change in the Northern Great Plains as recorded in the geochemistry of sediments in Pickerel Lake, South Dakota. Quat Int 67: 5–20

Delworth TL, Knutson TR (2000) Simulation of early 20th century global warming. Science 287: 2246–2250

Delworth TL, Mann ME (2000) Observed and simulated multidecadal variability in the Northern Hemisphere. Clim Dyn 16: 661–676

deMenocal P (2001) Cultural responses to climate change during the late Holocene. Science 292: 667–673

deMenocal P, Ortiz J, Guilderson T, Sarnthein M (2000) Coherent high- and low-latitude climate variability during the Holocene warm period. Science 288: 2198–2202

Domack EW, Mayewski PA (1999) Bi-polar ocean linkages Evidence from late-Holocene Antarctic marine and Greenland ice-core records. Holocene 9 247–251

Domack E and 5 others plus ODP Leg 178 Team (2001) Chronology of the Palmer Deep site, Antarctic Peninsula: A Holocene palaeoenvironmental reference for the circum-Antarctic. Holocene 11: 1–9

Dommenget D, Latif M (2002) A cautionary note on the interpretation of EOFs. J Clim 15: 216–225

Doran PT and 12 others (2002) Antarctic climate cooling and terrestrial ecosystem response. Nature 415: 517–520; See also exchanges between Turner et al. and Walsh, Doran et al. (2002) in Recent temperature trends in the Antarctic. Nature 418: 291–292

Druffel EM (1982) Banded corals: Changes in oceanic carbon-14 during the Little Ice Age. Science 218: 13–19

Dyurgerov MB, Meier MF (2000) Twentieth century climate change: Evidence from small glaciers. Proc Natl Acad Sci USA 97: 1406–1411

Ely LL, Enzel Y, Baker VR, Cayan DR (1993) A 5000-year record of extreme floods and climate change in the Southwestern United States. Science 262: 410–412

Emery WJ, Castro S, Wick GA, Schläussel P, Donlon G (2001a) Estimating sea surface temperature from infrared satellite and in situ temperature data. Bull Am Meteor Soc 82: 2773–2785

Emery WJ, Baldwin DJ, Schläussel P, Reynolds RW (2001b) Accuracy of in situ sea surface temperatures used to calibrate infrared satellite measurements. J Geophys Res 106: 2387–2405

Esper J, Cook ER, Schweingruber FH (2002a) Low-frequency signals in long tree-ring chronologies for reconstructing past temperature variability. Science 295: 2250–2253

Esper J, Schweingruber FH, Winiger M (2002b) 1300 years of climate history for Western Central Asia inferred from tree-rings Holocene 12: 267–277

Etheridge DM and 5 others (1996) Natural and anthropogenic changes in atmospheric CO_2 over the last 1000 years from air in Antarctic ice and firn. J Geophys Res 101: 4115–4128

Evans DJA (2000) Ice in a greenhouse: The impacts of global warming on world glacier cover. Trans Inst Engineers and Shipbuilders in Scotland 143: 11–19

Evans MN, Kaplan A, Cane MA (2002) Pacific sea surface temperature field reconstruction from coral $\delta^{18}O$ data using reduced space objective analysis. Paleocean 17: 10.1029/2000PA000590

Evans MN, Kaplan A, Cane MA, Villalba R (2000) Globality and optimality in climate field reconstructions from proxy data. In Markgraf V (ed) Inter-hemispheric Climate Linkages, Academic Press, p. 53–72

Feng X (1999) Trends in intrinsic water-use efficiency of natural trees for the past 100–200 years. Geochimica et Cosmochimica Acta 63: 1891–1903

Feng X, Epstein S (1994) Climatic implications of an 8000-year hydrogen isotope time series from Bristlecone Pine trees. Science 265: 1079–1081

Filippi ML, Lambert P, Hunziker J, Käubler B, Bernasconi S (1999) Climatic and anthropogenic influence on the stable isotope record from bulk carbonates and ostracodes in Lake Neuchatel, Switzerland, during the last two millennia. J Paleolim 21: 19–34

Folland CK and 10 others (2001) Global temperature change and its uncertainties since 1861. Geophys Res Lett 28: 2621–2624

Forster G (2000) A Voyage Round the World, 2 vols. edited by N. Thomas and O. Berghof, University of Hawaii Press

Fritz SC, Ito E, Yu Z, Laird KR, Engstrom DR (2000) Hydrologic variation in the Northern Great Plains during the last two millennia. Quat Res 53: 175–184

St. George S, Nielsen E (2002) Hydroclimatic change in Southern Manitoba since A.D. 1409 inferred from tree rings. Quat Res 58: 103–111

Goody R, Anderson J, North G (1998) Testing climate models: An approach. Bull Am Meteor Soc 79: 2541–2549

Graumlich LJ (1993) A 1000-year record of temperature and precipitation in the Sierra Nevada. Quat Res 39: 249–255

Graybill DA, Idso SB (1993) Detecting the aerial fertilization effect of atmospheric CO_2 enrichment in tree-ring chronologies. Glob BioGeoChem Cyc 7: 81–95

Grissino-Mayer HD (1996) A 2129-year reconstruction of precipitation for northwestern New Mexico, USA. In Dean JS, Meko DM, Swetnam TW (eds) Tree rings, environment and humanity, Radiocarbon (U Arizona Press), p. 191–204

Grove AT (2001) The "Little Ice Age" and its geomorphological consequences in Mediterranean

Europe. Clim Change 48: 121–136

Grove JM (1996) The century time-scale. In Driver TS, Chapman GP (eds) Time-scales and environmental change, Routledge, p. 39–87

Grove JM (2001a) The initiation of the Little Ice Age in regions round the North Atlantic. Clim Change 48: 53–82

Grove JM (2001b) The onset of the Little Ice Age. In Jones PD, Ogilvie AEJ, Davies TD, Briffa KR (eds) History and Climate – Memories of the Future?, Kluwer, p. 153–185

Grove JM, Conterio A (1995) The climate of Crete in the sixteenth and seventeenth centuries. Clim Change 30: 223–247

Grove JM, Switsur R (1994) Glacial geological evidence for the medieval warm period. Clim Change 26: 143–169

Grove RH (1998) Global impact of the 1789–93 El Nino. Nature 393: 318–319

Hadly EA (1996) Influence of late-Holocene climate on Northern Rocky Mountain mammals. Quat Res 46: 289–310

Häkkinen S (2000) Decadal air-sea interaction in the North Atlantic based on observations and modelling results. J Clim 13: 1195–1219

Hansen J, Ruedy R, Glascoe J, Sato M (1999) GISS analysis of surface temperature change. J Geophys Res 104: 30997–31022

Harris RN, Chapman DS (2001) Mid-latitude (30°–60°N) climatic warming inferred by combining borehole temperatures with surface air temperatures. Geophys Res Lett 28: 747–750

Harrison DE, Larkin NK (1998) Seasonal U.S. temperature and precipitation anomalies associated with El Niño: Historical results and comparison with 1997–98. Geophys Res Lett 25: 3959–3962

Haug GH, Hughen KA, Sigman DM, Peterson LC, Röhk U (2001) Southward migration of the intertropical convergence zone through the Holocene. Science 293: 1304–1308

Heikkiläa M, Seppä H (2003) A 11,000 yr paleotemperature reconstruction from the southern boreal zone in Finland. Quat Sci Rev : in press

Henderson GM (2002) New oceanic proxies for paleoclimate. Earth and Planetary Science Letters 203: 1–13

Hendy EJ and 5 others (2002) Abrupt decrease in tropical Pacific sea surface salinity at end of Little Ice Age. Science 295: 1511–1514

Hiller A, Boettger T, Kremenetski C (2001) Mediaeval climatic warming recorded by radiocarbon date alpine tree-line shift on the Kola Peninsula, Russia. Holocene 11: 491–497

Hodell DA, Brenner M, Curtis JH, Guilderson T (2001) Solar forcing of drought frequency in the Maya lowlands. Science 292: 1367–1370

Holzhauser H (1997) Fluctuations of the Grosser Aletsch Glacier and the Groner Glacier during the last 3200 years: New results. In Frenzel B, Boulton GS, Gläser B, Huckriede U (eds) Glacier fluctuations during the Holocene, Akademie der Wissenschaften und der Literatur, p. 35–58

Hong YT and 8 others (2000) Response of climate to solar forcing recorded in a 6000-year $\delta^{18}O$ time-series of Chinese peat cellulose. Holocene 10: 1–7

Hormes A, Muller BU, Schluchter C (2001) The Alps with little ice: Evidence for eight Holocene phases of reduced glacier extent in the Central Swiss Alps. Holocene 11: 255–265

Hsu HH, Moura AD (2001) Meeting summary: Workshop on the impacts of the 1997–99 ENSO. Bull Am Meteor Soc 82: 305–312

Hu FS, Ito E, Brown TA, Curry BB, Engstrom DR (2001) Pronounced climatic variations in Alaska during the last two millennia. Proc Natl Acad Sci USA 98: 10552–10556

Huang S, Pollack HN (2002) Borehole temperature observations compared to synthetic subsurface temperatures generated by proxy climate reconstructions (abstract). EOS Trans AGU suppl 83: p. F 918

Huang S, Pollack HN, Shen PY (1997) Late Quaternary temperature changes seen in the world-wide continental heat flow measurements. Geophys Res Lett 24: 1947–1950

Huang S, Shen PY, Pollack HN (1996) Deriving century-long trends of surface temperature change from borehole temperatures. Geophys Res Lett 23: 257–260

Huang S, Pollack HN, Shen PY (2000) Temperature trends over the past five centuries reconstructed from borehole temperatures. Nature 403: 756–758

Huffman TN (1996) Archaeological evidence for climatic change during the last 2000 years in Southern Africa. Quat Int 33 55–60

Hughes MK, Diaz HF (1994) Was there a 'Medieval Warm Period', if so, where and when? Clim Change 26: 109–142

Hurrell JW, Trenberth KE (1999) Global sea surface temperature analyses: Multiple problems and their implications for climate analysis, modelling, and reanalysis. Bull Am Meteor Soc 80: 2661–2678

Huybrechts P, de Wolde J (1999) The dynamic response of the Greenland and Antarctic ice sheets to multiple-century climatic warming. J Clim 12: 2169–2188

Idso SB (1989) Letter to the editor – A problem for paleoclimatology? Quat Res 31: 433–434

Idso SB (1998) CO_2-induced global warming: A skeptic's view of potential climate change Clim Res 10: 69–82

Ingram MJ, Underhill DJ, Wigley TML (1978) Historical climatology. Nature 276: 329–334

Iriondo M (1999) Climatic changes in the South American plains: Records of a continent-scale oscillation. Quat Int 57–58: 93–112

Jacoby GC, D'Arrigo RD (1995) Tree-ring width and density evidence for climatic and potential forest change in Alaska Glob Biogeochem Cyc 9: 227–234

Jacoby GC, D'Arrigo RD, Davaajamts T (1996) Mongolian tree rings and 20th-century warming. Science 273 771–773

Jacoby GC and 5 others (2000) Long-term temperature trends and tree growth in the Taymir region of Northern Siberia. Quat Res 53: 312–318

Jennings AE, Weiner NJ (1996) Environmental change in eastern Greenland during the last 1300 years: Evidence from foraminifera and lithofacies in Nansen Fjord, 68°N. Holocene 6: 171–191

Jenny B and 6 others (2002) Moisture changes and fluctuations of the Westerlies in Mediterranean Central Chile during the last 2000 years: The Laguna Aculeo record (33°50' S). Quat Int 87: 3–18

Johnson TC, Barry S, Chan Y, Wilkinson P (2001) Decadal record of climate variability spanning the past 700 yr in the Southern tropics of East Africa. Geology 29: 83–86

Jones PD, Briffa KR, Barnett TP, Tett SFB (1998) High-resolution palaeoclimatic records for the last millennium: Interpretation, integration and comparison with General Circulation Model control-run temperatures. Holocene 8: 455–471

Jones PD, Moberg A (2003) Hemispheric and large-scale surface air temperature variations: An extensive revision and an update to 2001. J Clim 16: 206–223

Jones PD, New M, Parker DE, Martin S, Rigor IG (1999) Surface air temperatures and its changes over the past 150 years. Rev Geophys 37: 173–199

Jones PD, Osborn TJ, Briffa KR (2001) The evolution of climate over the last millennium. Science 292: 662–667

Karlén W (1998) Climate variations and the enhanced greenhouse effect. Ambio 27: 270–274

Karlén W (2001) Global temperature forced by solar irradiation and greenhouse gases? Ambio 30: 349–350

Karlöf L and 13 others (2000) A 1500 year record of accumulation at Amundsenisen western Dronning Maud Land, Antarctica, derived from electrical and radioactive measurements on a 120 m ice core. J Geophys Res 105: 12471–12483

Kaser G (1999) A review of the modern fluctuations of tropical glaciers. Glob Planet Change 22: 93–103

Keigwin LD (1996) The little ice age and medieval warm period in the Sargasso Sea. Science 274: 1504–1508

Keith DW, Anderson JG (2001) Accurate spectrally resolved infrared radiance observation from space: Implications for the detection of decade-to-century-scale climatic change. J Clim 14: 979–990

Kennett JP, Hodell DA (1995) Stability or instability of Antarctic ice sheets during the warm climates of the Pliocene ? GSA Today 5: 1–22 (about 5 pp.)

Khim BK, Yoon HI, Kang CY, Bahk JJ (2002) Unstable climate oscillations during the late Holocene in the Eastern Bransfield Basin, Antarctic Peninsula. Quat Res 58: 234–245

Kiladis GN, Diaz HF (1986) An analysis of the 1877–78 ENSO episode and comparison with 1982–83. Mon Wea Rev 114: 1035–1047

Kitagawa H, Matsumoto E (1995) Climatic implications of δ^{13}C variations in a Japanese cedar (*Cryptomeria japonica*) during the last two millennia. Geophys Res Lett 22: 2155–2158

Knapp PA, Soulé PT, Grissino-Mayer HD (2001) Post-drought growth responses of western

juniper (*Juniperus occidentalis* var. *occidentalis*) in central Oregon. Geophys Res Lett 28: 2657–2660

Knappenberger PC, Michaels PJ, Davis RE (2001) Nature of observed temperature changes across the United States during the 20th century. Clim Res 17: 45–53

Koerner RM (1977) Devon Island ice cap: Core stratigraphy and paleoclimate. Science 196: 15–18

Koerner RM, Fisher DA (1990) A record of Holocene summer climate from a Canadian high-Arctic ice core. Nature 343: 630–631

Korhola A, Weckström J, Holmström L, Erasto P (2000) A quantitative Holocene climatic record from diatoms in Northern Fennoscandia Quat Res 54: 284–294

Kreutz KJ and 5 others (1997) Bipolar changes in atmospheric circulation during the Little Ice Age. Science 277: 1294–1296

Krishnamurti TN and 7 others (2000) Coupled atmosphere-ocean modelling of the El Niño of 1997–98. J Clim 13: 2428–2459

Ku TL, Li HC (1998) Speleothems as high-resolution paleoenvironment archives: Records from northeastern China. Proc Indian Acad Sci (Earth Planet Sci) 107: 321–330

Kuhnert H, Pätzold J, Schnetger B, Wefer G (2002) Sea-surface temperature variability in the 16th century at Bermuda inferred from coral records. Palaeogeogr Palaeoclim Palaeoecol 179: 159–171

Kumar A, Wang WQ, Hoerling MP, Leetma A, Ji M (2001) The sustained North American warming of 1997 and 1998. J Clim 14: 345–353

Laird KR, Fritz SC, Maasch KA, Cumming BF (1996) Greater drought intensity and frequency before AD 1200 in the Northern Great Plains, USA. Nature 384: 552–554

Lamb HH (1963) On the nature of certain climatic epochs which differed from the modern (1900–39) normal. In Proceedings of the WMO-UNESCO Symposium on Climate Change, Rome, 1961. UNESCO Arid Zone Res. 20, p. 125–150

Lamb HH (1965) The early medieval warm epoch and its sequel. Palaeogeogr Palaeoclim Palaeoecol 1: 13–37

Lamb HH (1977) Climate – Past, Present and Future Vol. 2 Climatic History and Future, Methuen

Lamb HH (1982) Climate history and the modern world, Methuen

Lamb HH (1997a) Climate history and the modern world, 2nd Edition, Routledge

Lamb H (1997b) Through all the changing scenes of life: A meteorologist's tale, Taverner

Lamoureux SF, Bradley RS (1996) A late Holocene varved sediment record of environmental change from northern Ellesmere Island, Canada. J Paleolim 16: 239–255

Lara A, Villalba R (1993) A 3260-year temperature record from *Fitzroya cupressoides* tree rings in southern South America. Science 260: 1104–1106

Lau NC, Nath MJ (2001) Impact of ENSO and SST variabiluty in the North Pacific and North

Atlantic: Seasonal dependence and role of extratropical sea-air coupling. J Clim 14: 2846–2866

Lau KM, Weng HY (2001) Coherent modes of global SST and summer rainfall over China: An assessment of the regional impacts of the 1997–98 El Niño. J Clim 14: 1294–1308

Lawton RO, Nair US, Pielke Sr. RA, Welch RM (2001) Climatic impact of tropical lowland deforestation on nearby montane cloud forests. Science 294: 584–587

Leroux M (1993) The Mobile Polar High: A new concept explaining present mechanisms of meridional air-mass and energy exchanges and global propagation of palaeoclimatic changes. Glob Planet Change 7: 69–93

Lewis TJ, Wang K (1998) Geothermal evidence for deforestation induced warming: Implications for the climatic impact of land development. Geophys Res Lett 25: 535–538

Li HC, Bischoff JL, Ku TL, Lund SP, Stot LD (2000) Climate variability in East-Central California during the past 1000 years reflected by high-resolution geochemical and isotopic records from Owens Lake sediments. Quat Res 54: 189–197

Li HC, Ku TL, Stot LD, Chen W (1998) Applications of interannual-resolution stable isotope records of speleothem: Climatic changes in Beijing and Tianjin, China during the past 500 years – the $\delta^{18}O$ record. Sci in China (Series D) 41: 362–368

Liu K, Yao Z, Thompson LG (1998) A pollen record of Holocene climatic changes from Dunde ice cap, Qinghai-Tibetan Plateau. Geology 26: 135–138

Lotter AF and 5 others (2000) Younger Dryas Alller d summer temperatures at Gerzensee (Switzerland) inferred from fossil pollen and cladoceran assemblages. Palaeogeogr Palaeoclim Palaeoecol 159: 349–361

Luckman BH (2000) The Little Ice Age in the Canadian Rockies. Geomorph 32: 357–384

Luckman BH, Briffa KR, Jones PD, Schweingruber FH (1997) Tree-ring based reconstruction of summer temperatures at Columbia Icefield, Alberta, Canada, AD 1073–1983. Holocene 7: 375–389

Luterbacher J and 33 others (2000) Monthly mean pressure reconstruction for the late Maunder Minimum period (AD 1675–1715). Int J Clim 20: 1049–1066

Madole RF (1994) Stratigraphic evidence of desertification in the west-central Great Plains within the past 1000 yr. Geology 22: 483–486

Manley G (1974) Central England temperatures: Monthly means 1659 to 1973. Quart J Roy Meteorol Soc 100: 389–405

Mann ME (2001a) Climate during the past millennium. Weather 56: 91–102

Mann ME (2001b) Little Ice Age. In MacCraken MC, Perry JS (eds) Encyclopedia of Global Environmental Change: Volume 1 – The Earth System: Physical and Chemical Dimensions of Global Environmental Change, John Wiley and Sons Ltd., p. 504–509

Mann ME (2002) The value of multiple proxies. Science 297: 1481–1482 Mann ME, Bradley RS, Hughes MK (1998) Global-scale temperature patterns and climate forcing over the past six centuries. Nature 392: 779–787

Mann ME, Bradley RS, Hughes MK (1999) Northern Hemisphere temperatures during the past

millennium· Inferences, uncertainties, and limitations. Geophys Res Lett 26: 759–762

Mann ME, Hughes MK (2002) Tree-ring chronologies and climate variability. Science 296: 848

Mann ME and 6 others (2000a) Global temperature patterns in past centuries: An interactive presentation. Earth Interact 4 (Paper No. 4): 1–29 [available online http://Earthinteraction.org]

Mann ME, Bradley RS, Hughes MK (2000b) Long-term variability in the El Nino Southern Oscillation and associated teleconnections. In Diaz HF, Markgraf V (eds) El Nino and the Southern Oscillation: Multiscale Variability and Global and Regional Impact, Cambridge Univ Press, p. 357–412

Mann ME, Rutherford S, Bradley RS, Hughes MK, Keimig FT (2003) Optimal surface temperature reconstructions using terrestrial borehole data. J Geophys Res: in press

Marshall GJ, Lagun V, Lachlan-Cope TA (2002) Changes in Antarctic Peninsula tropospheric temperatures from 1956 to 1999: A synthesis of observations and reanalysis data. Int J Clim 22: 291–310

Marshall J, Johnson H, Goodman J (2001) A study of the interaction of the North Atlantic oscillation with ocean circulation. J Clim 14: 1399–1421

Mayewski PA, White F (2002) The last thousand years of climate change. In The Ice Chronicles, University Press of New England, p. 131–145

McDermott F, Mattey DP, Hawkesworth C (2001) Centennial-scale Holocene climate variability revealed by a high-resolution speleothem $\delta^{18}O$ record from SW Ireland. Science 294: 1328–1331

Meeker LD, Mayewski PA, Grootes PM, Alley RB, Bond GC (2001) Comment on "On sharp spectral lines in the climate record and the millenial peak" by Carl Wunsch. Paleocean 16: 544–547

Meese DA and 8 others (1994) The accumulation record from the GISP2 core as an indicator of the climate change throughout the Holocene. Science 266: 1680–1682

Mercer JH (1978) West Antarctic ice sheet and CO_2 greenhouse effect: A threat of disaster. Nature 271: 321–325

Messerli B, Grosjean M, Hofer T, Nunez, L, Pfister C. (2000) From nature-dominated to human-dominated environmental changes. Quat Sci Rev 19: 459–479

Morgan VI (1985) An oxygen isotope climate record from the Law Dome, Antarctica. Clim Change 7: 415–426

Mosley-Thompson E, Thompson LG (1982) Nine centuries of microparticles deposition at the South Pole. Quat Res 17: 1–13

Mosley-Thompson E (1995) Paleoenvironmental conditions in Antarctica since A.D. 1500: Ice core evidence. In Bradley RS, Jones PD (eds) Climate Since A.D. 1500, Routledge, p. 572–591

Naurzbaev MM, Vaganov EA (2000) Variation of early summer and annual temperature in east Taymir and Putorian (Siberia) over the last two millennia. J Geophys Res 105: 7317–7326

Niggemann S, Mangini A, Richter DK, Wurth G (2003) A paleoclimate record of the last 17,600 years in stalagmites from the B7 cave, Sauerland, Germany. Quat Sci Rev: in press

Nyberg J, Malmgren BA, Kuijpers A, Winter A (2002) A centennial-scale variability of tropical North Atlantic surface hydrography during the late Holocene. Palaeogeogr Palaeoclim Palaeoecol 183: 25–41

Ogilvie AEJ (1984) The past climate and sea-ice record from Iceland, Part 1: Data to A.D. 1780. Clim Change 6: 131–152

Ogilvie AEJ, Barlow LK, Jennings AE (2000) North Atlantic climate c. A.D. 1000: Millennial reflections on the Viking discoveries of Iceland, Greenland and North America. Weather 55: 34–45

Ogilvie A, Farmer G (1997) Documenting the medieval climate. In Hulme M, Barrow E (eds) Climate of the British Isles – Present, Past and Future, Routledge, p. 112–133

Ogilvie AEJ, Jónsson T (2001) "Little Ice Age" research: A perspective from Iceland. Clim Change 48: 9–52

Ou HW (2001) Possible bounds on the Earth's surface temperature: From the perspective of a conceptual global-mean model. J Clim 14: 2976–2988

Overpeck J and 17 others (1997) Arctic environmental change of the last four centuries. Science 278: 1251–1256

Parker DE, Folland CK, Jackson M (1995) Marine surface temperature: Observed variations and data requirements. Clim Change 31: 559–600

Paulsen DE, Li HC, Ku TL (2003) Climate variability in central China over the last 1270 years revealed by high-resolution stalagmite records. Quat Sci Rev: in press

Pederson JL (2000) Holocene paleolakes of Lake Canyon, Colorado Plateau: Paleoclimate and landscape response from sedimentology and allostratigraphy. Geol Soc Amer Bull 112: 147–158

Perry CA, Hsu KJ (2000) Geophysical, archaeological, and historical evidence support a solar-output model for climate change. Proc Natl Acad Sci USA 97: 12433–12438

Petersen KL (1994) A warm and wet little climatic optimum and a cold and dry little ice age in the Southern Rocky Mountains, USA. Clim Change 26: 243–269

Pfister C (1995) Monthly temperature and precipitation in central Europe 1525–1979: Quantifying documentary evidence on weather and its effects. In Bradley RS, Jones PD (eds) Climate Since A.D. 1500, Routledge, p. 118–142

Pfister C, Luterbacher J, Schwarz-Zanetti G, Wegmann M (1998) Winter air temperature variations in western Europe during the early and high Middle ages (AD 750–1300). Holocene 8: 535–552

Pielke Jr RA, (1998) Rethinking the role of adaption in climate policy. Glob Environ Change 8: 159–170

Pielke Sr RA, (2001) Earth system modelling – An integrated assessment tool for environmental studies. In Matsuno T, Kida H (eds) Present and future of modelling global environmental change: Toward integrated modelling, TERRAPUB Tokyo, p. 311–337

Pisias NG (1978) Paleoceanography of the Santa Barbara Basin during the last 8000 years. Quat Res 10: 366–384

Porter SC (1981) Glaciological evidence of Holocene climatic change. In Wigley TML, Ingram MJ, Farmer G (eds) Climate and History – Studies in past climates and their impact on Man, Cambridge University Press, p. 82–110

Porter SC (1986) Pattern and forcing of the Northern hemisphere glacier variations during the last millennium. Quat Res 26: 27–48

Proctor CJ, Baker A, Barnes WL, Gilmour MA (2000) A thousand year speleothem proxy record of North Atlantic climate from Scotland. Clim Dyn 16: 815–820

Qian W, Zhu Y (2002) Little Ice Age climate near Beijing, China, inferred from historical and stalagmite records. Quat Res 57: 109–119

Ramesh R (1993) First evidence for Little Ice Age and Medieval Warming in India (abstract). EOS Trans AGU suppl 74: 118

Räisäanen J (2001) CO_2-induced climate change in CMIP2 experiments: Quantification of agreement and role of internal variability. J Clim 14: 2088–2104

Ren G (1998) Pollen evidence for increased summer rainfall in the Medieval warm period at Maili, Northeast China. Geophys Res Lett 25: 1931–1934

Robinson WA, Ruedy R, Hansen J (2002) General circulation model simulations of recent cooling in the east-central United States. J Geophys Res 107 (D24), 4748, doi:10.1029/2001JD001577

Rodrigo FS, Esteban-Parra MJ, Pozo-Vázquez D, Castro-Díez Y (2000) Rainfall variability in Southern Spain on decadal to centennial time scales. Int J Clim 20: 721–732

Rott H, Skvarca P, Nagler T (1996) Rapid collapse of Northern Larsen ice shelf, Antarctica. Science 271: 788–792

Sakamoto N, Masuda K (2002) Significance of the sampling errors due to changing land-sea distribution in the Phanerozoic paleo-climate reconstruction. J Meteor Soc Japan 80: 531–537

Schleser GH, Helle G, Lucke A, Vos H (1999) Isotope signals as climate proxies: The role of transfer functions in the study of terrestrial archives. Quat Sci Rev 18: 927–943

Schmutz C, Luterbacher J, Gyalistras D, Xoplaki E, Wanner H (2000) Can we trust proxy-based NAO index reconstructions? Geophys Res Lett 27: 1135–1138

Schneider T (2001) Analysis of incomplete climate data: Estimation of mean values and covariance matrices and imputation of missing values. J Clim 14: 853–871

Schuster PF, White DE, Naftz DL, Cecil LD (2000) Chronological refinement of an ice core record at Upper Fremont Glacier in south central North America. J Geophys Res 105: 7317–7326

Scuderi LA (1993) A 2000-year tree ring record of annual temperatures in the Sierra Nevada mountains. Science 259: 1433–1436

Seager R and 6 others (2000) Causes of Atlantic ocean climate variability between 1958 and 1998. J Clim 13: 2845–2862

Seppä H, Birks HJB (2002) Holocene climate reconstructions from the Fennoscandian tree-line area based on pollen data from Toskaljavri. Quat Res 57: 191–199

Serre-Bachet F (1994) Middle ages temperature reconstructions in Europe, a focus on Northeastern Italy. Clim Change 26: 213–240

Shabalova MV, Weber SL (1998) Seasonality of low-frequency variability in early- instrumental European temperatures. Geophys Res Lett 25: 3859–3862

Shi Y, Yao T, Yang B (1999) Decadal climatic variations recorded in Guliya ice core and comparison with the historical documentary data from East China during the last 2000 years. Sci in China Ser D (suppl) 42: 91–100

Skinner WR, Majorowicz JA (1999) Regional climatic warming and associated twentieth century land-cover changes in north-western North America. Clim Res 12: 39–52

Slonosky VC, Yiou P (2001) The North Atlantic Oscillation and its relationship with near surface temperature. Geophys Res Lett 28: 807–810

Smith RC and 10 others (1999) Marine ecosystem sensitivity to climate change. Bioscience 49: 393–404

Song J (2000) Changes in dryness/wetness in China during the last 529 years. Int J Clim 20: 1003–1015

Stahle DW, Cleaveland MK, Hehr JG (1988) North Carolina climate changes reconstructed from tree rings: A.D. 372 to 1985. Science 240: 1517–1519

Stenni B and 6 others (2002) Eight centuries of volcanic signal and climate change at Talos Dome (East Antarctica). J Geophys Res 107: doi:10.1029/2000JD000317

Stine S (1994) Extreme and persistent drought in California and Patagonia during medieval time. Nature 369: 546–549

Stine S (1998) Medieval climatic anomaly in the Americas. In Issar AS, Brown N (eds) Water, Environment and Society in Times of Climatic Change, Kluwer Academic Press, p. 43–67

Stott L (2002) SST variability in the Western Pacific Warm Pool during the past 2000 years (abstract). EOS Trans AGU suppl 83: p. F 915

Stott L, Poulsen, C, Lund S, Thunell, R (2002) Super ENSO and global climate oscillations at millennial time scales. Science 297: 222–226 (plus erratum posted October 25, 2002)

Stuiver M, Grootes PM, Braziunas TF (1995) The GISP2 $\delta^{18}O$ climate record of the past 16,500 years and the role of the Sun, ocean and volcanoes. Quat Res 44: 341–354

Sugden DE (1996) The East Antarctic ice sheet: Unstable ice or unstable ideas? Trans Inst Br Geogr NS 21: 443–454

Swetnam TW (1993) Fire history and climate change in Giant Sequoia groves. Science 262: 885–889

Tagami Y (1993) Climate change reconstructed from historical data in Japan. In Proceedings of International Symposium on Global Change by IGBP, International Geosphere-Biosphere Programme-IGBP, p. 720–729

Tagami Y (1996) Some remarks on the climate in the Medieval Warm Period of Japan. In Mikami T, Matsumoto E, Ohta S, Sweda T (eds) Paleoclimate and environmental variability in Austral-Asian transect during the past 2000 years (IGBP Proceedings), International Geosphere-

Biosphere Programme-IGBP, p. 115–119

Tarussov A (1995) The Arctic from Svalbard to Severnaya Zemlya: Climatic reconstructions from ice cores. In Bradley RS, Jones PD (eds) Climate Since A.D. 1500, Routledge, p. 505–516

Thomas RH, Sanderson TJO, Rose KE (1979) Effect of climatic warming on the West Antarctic ice sheet. Nature 277: 355–358

Thompson DWJ, Baldwin MP, Wallace JM (2002) Stratospheric connection to northern hemisphere wintertime weather: Implications for prediction. J Clim 15: 1421–1428

Thompson LG (2000) Ice core evidence for climate change in the Tropics: Implications for our future. Quat Sci Rev 19: 19–35

Thompson LG, Mosley-Thompson E, Dansgaard W, Grootes PM (1986) The little ice age as recorded in the stratigraphy of the tropical Quelccaya ice cap. Science 234: 361–364

Thompson LG and 9 others (1989) Holocene-Late Pleistocene climatic ice core records from Qinghai-Tibetan plateau. Science 246: 474–477

Thompson LG and 7 others (1994) Climate since AD 1510 on Dyer Plateau, Antarctic Peninsula: Evidence for recent climate change. Ann Glaciology 20: 420–426

Thompson LG and 6 others (1995) A 1000 year climate ice-core record from the Guliya ice cap, China: Its relationship to global climate variability. Ann Glaciology 21: 175–181

Thompson LG and 5 others (2000) A high-resolution millennial record of the South Asian Monsoon from Himalayan ice cores. Science 289: 1916–1919

Till C, Guiot J (1990) Reconstruction of precipitation in Morocco since 1100 A.D. based on *Cedrus atlantica* tree-ring widths. Quat Res 33: 337–351

Tomita T, Wang B, Yasunari T, Nakamura H (2001) Global patterns of decadal scale variability observed in sea surface temperature and lower-tropospheric circulation fields. J Geophys Res 106 26805–26815

Tong G and 7 others (1996) Environmental changes at the top of Taibai Mountain, Qinling since 1ka ago. Mar Geol Quat Geol 16: 95–104 (In Chinese but English Abstract available)

Tyson PD, Karlén W, Holmgren K, Heiss GA (2000) The Little Ice Age and medieval warming in South Africa. South African J Sci 96: 121–126

Tyson PD, Lindesay JA (1992) The climate of the last 2000 years in southern Africa. Holocene 2: 271–278

Ueda H, Matsumoto J (2000) A Possible Triggering Process of East-West Asymmetric Anomalies over the Indian Ocean in Relation to 1997/98 El Nino. J Meteor Soc Japan 78: 803–818

Valero-Garcés BL, Delgado-Huertas A, Ratto N, Navas A, Edwards L (2000) Paleohydrology of Andean saline lakes from sedimentological and isotopic records, Northwestern Argentina. J Paleolimnology 24: 343–359

van Engelen AFV, Buisman J, Ijnsen F (2001) A millennium of weather, winds and water in the low countries. In Jones PD, Ogilvie AEJ, Davies TD Briffa KR (eds) History and Climate – Memories of the Future?, Kluwer, p. 101–124

Vaughan DG, Doake CSM (1996) Recent atmospheric warming and retreat of ice shelves on the Antarctic Peninsula. Nature 379: 328–331

Verardo S, Cronin T, Willard D, Kerhin R (1998) Climate and environmental changes recorded by Chesapeake Bay Dinoflagellate cysts in the last millennium. Geol Soc Amer Abs 74: 284

Verschuren D, Laird KR, Cumming BF (2000) Rainfall and drought in equatorial east Africa during the past 1,100 years. Nature 403: 410–414

Villalba R (1990) Climatic fluctuations in Northern Patagonia during the last 1000 years as inferred from tree-ring records. Quat Res 34: 346–360

Villalba R (1994) Tree-ring and glacial evidence for the medieval warm epoch and the little ice age in southern South America. Clim Change 26: 183–197

Visser K, Thunell R, Stott L (2003) Magnitude and timing of temperature change in the Indo-Pacific warm pool during deglaciation. Nature 421: 152–155

von Storch JS, Muller P, Bauer E (2001) Climate variability in millennium integrations with coupled atmosphere-ocean GCMs: A spectral view. Clim Dyn 17: 375–389

Williams PW, Marshall A, Ford DC, Jenkinson AV (1999) Palaeoclimatic interpretation of stable isotope data from Holocene speleothems of the Waitomo district, North Island, New Zealand. Holocene 9: 649–657

Wilson AT, Hendy CH, Reynolds CP (1979) Short-term climate change and New Zealand temperatures during the last millennium. Nature 279: 315–317

Winkler S (2000) The 'Little Ice Age' maximum in the Southern Alps, New Zealand: Preliminary results at Mueller Glacier. Holocene 10: 643–647

Winter A, Ishioroshi H, Watanabe T, Oba T, Christy J (2000) Carribean sea surface temperatures: Two-to-three degrees cooler than present during the Little Ice Age. Geophys Res Lett 27: 3365–3368

Woodhouse CA, Overpeck JT (1998) 2000 years of drought variability in the Central United States. Bull Am Meteor Soc 79: 2693–2714

Wunsch C (1999) The interpretation of short climate records, with comments on the North Atlantic and Southern Oscillations. Bull Am Meteor Soc 80: 245–255 (See also important research exchanges in Bull Am Meteor Soc 80: 2721–2728)

Wunsch C (2000) On sharp spectral lines in the climate record and the millenial peak. Paleocean 15: 417–424

Wunsch C (2001) Reply to Meeker et al. (2001). Paleocean 16: 548

Yadav RR, Singh J (2002) Tree-ring-based spring temperature patterns over the past four centuries in Western Himalaya. Quat Res 57: 299–305

Yang B, Braeuning A, Johnson KR, Shi Y (2002) General characteristics of temperature variation in China during the last two millennia. Geophys Res Lett 29: 10.1029/2001GL014485, 38-1–38-4

Yi S, Saito Y, Oshima H, Zhou Y, Wei H (2003) Holocene environmental history inferred from pollen assemblages in the Huanghe (Yellow River) delta, China: Climatic change and human

impact. Quat Sci Rev: in press

Yu Z, Ito E (1999) Possible solar forcing of century-scale drought frequency in the northern Great Plains. Geology 27: 262–266

Zhang D (1994) Evidence for the existence of the medieval warm period in China. Clim Change 26: 289–297

Zorita E, Gonzalez-Ruoco F, Legutke S (2003) Testing the Mann et al. (1998) approach to paleoclimate reconstructions in the context of a 1000 year simulation with the ECHO-G coupled climate model. J Clim: in press

Table 1: A full list of paleoclimatic proxies that have sufficient length of continuous records to entertain the three specific questions:

(1) Is there an objectively discernible climatic anomaly during the Little Ice Age interval (1300-1900 A.D.) in this proxy record?
(2) Is there an objectively discernible climatic anomaly during the Medieval Warm Period (800-1300 A.D.) in this proxy record?
(3) Is there an objectively discernible climatic anomaly within the 20th century that may validly be considered the most extreme (the warmest, if such information is available) period in the record?

We list the location, type of proxies, reference and the logical answers to the three specific questions posed in this study.

Location	Lat.	Long.	Type	Reference	Ans. (1)	Ans. (2)	Ans. (3)
World-wide	–	–	Mp	Mann et al. 99	yes	no	yes*
Arctic-wide	–	–	Mp	Overpeck et al. 97	yes	–	yes*
World-wide	–	–	Mp	Crowley & Lowery 00	yes	no	yes*
World-wide	–	–	Mp	Jones et al. 98	yes	no	yes*
World-wide	–	–	T	Briffa 00	yes	no	yes*
World-wide	–	–	T	Briffa et al. 01	yes	–	yes*
World-wide	–	–	T	Jones et al. 01	yes	–	yes*
NH Mid-Latitude	–	–	T	Esper et al. 02a	yes	yes	no
World-wide	–	–	Mp	Lamb 77, 82	yes	yes	–
World-wide	–	–	G + Is	Porter 86	yes	yes	–
World-wide	–	–	G	Grove & Switsur 94	–	yes	–
World-wide	–	–	T+G+D	Hughes & Diaz 94	yes	no?+	no?+
World-wide	–	–	Mp	Grove 96	–	yes	–
World-wide	–	–	B	Huang et al. 97	yes	yes	no
World-wide	–	–	D	Perry & Hsu 00†	yes	yes	no
World-wide	–	–	D	deMenocal 01	yes	yes	–
China-wide	–	–	Mp	Yang et al. 02	yes	yes	–
Americas	–	–	Ts+Gm+Mp	Stine 98	–	yes	–

+ In reality, Hughes & Diaz concluded that "[o]ur review indicates that for some areas of the globe (for example, Scandinavia, China, the Sierra Nevada in California, the Canadian Rockies and Tasmania), temperatures, particularly in summer, appear to have been higher during some parts of this period than those that were to prevail until the most recent decades of the twentieth century. These regional episodes were not strongly synchronous. Evidence from other regions (for example, the Southeast United States, southern Europe along the Mediterranean, and parts of South America) indicates that the climate during that time was little different to that of later times, or that warming, if it occurred, was recorded at a later time than has been assumed... To the extend that glacial retreat is associated with warm summers, the glacial geology evidence would be consistent with a warmer period in A.D. 900-1250 than immediately before or for most of the following seven hundred years." We simply note that the main conclusion of Hughes & Diaz (1994) may be in actual agreement with the qualitative classification in our paper.

† We refer only to the documentary, historical and archaeological research results, rather than the solar-output model results, of this paper.

Table 1 (continued)

Location	Lat.	Long.	Type	Reference	Ans. (1)	Ans. (2)	Ans. (3)
N. Atlantic (Iceland)	63–66°N	14–24°W	Mp	Ogilvie et al. 00, Ogilvie and Jónsson 01	yes	yes	no
N. Atlantic (S. Greenland)	60–70°N	20–55°W	Mp	Ogilvie et al. 00	yes	yes	no
W. Europe	45–54°N	0–15°E	Mp	Pfister et al. 98	yes	yes	no
N. Atlantic (Europe)	35–70°N	25W–30°E	In+D	Luterbacher et al. 00	yes	–	–
Central England	52°N	2°E	In+D	Lamb 65, Manley 74	yes	yes	no
S. Spain	37.30°N	4.30°W	In+D	Rodrigo et al. 00	yes	–	no
Crete Is.	35.15°N	25.00°E	D	Grove & Conterio 95	yes	–	no
Mid–Russia	50–60°N	30–50°E	In+D	Borisenkov 95	yes	–	–
Czech Republic	48.5–51.2°N	12–19°E	In+B	Bodri & Cermák 99	yes	yes	yes?
S. USA	37–38°N	107.5–109.5°W	Pf+Cl+D	Petersen 94	yes	yes	–
E. China (GuangDong Prov.)	22–25°N	112–114.3°E	D	Chan & Shi 00	yes	–	–
E. China-wide	20–40°N	90–120°E	D	Song 00	yes	–	no?
Japan	30–40°N	125–145°E	D	Tagami 93, 96	yes	yes	no
S. Africa	22.2°S	28.38°E	Cl	Huffman 96	yes	yes	–
E. Greenland (Nansen Fjord)	68.3°N	29.7°W	Is	Jennings & Weiner 96	yes	yes	no
C. Greenland (Crête)	71.12°N	37.32°W	Is	Dansgaard et al. 75	yes	yes	no
C. Greenland (GRIP)	72.6°N	37.6°W	B	Dahl-Jensen et al. 98	yes	yes	no
S. Greenland (Dye 3)	65.2°N	43.8°W	B	Dahl-Jensen et al. 98	yes	yes	no
C. Greenland (GISP2)	72.58°N	38.5°W	Ic+Ml	Meese et al. 94	yes	yes	no
C. Greenland (GISP2)	72.58°N	38.5°W	Is	Stuiver et al. 95	yes	yes	no
Svalsbard	79°N	15°E	Ml	Tarussov 95	yes	yes	no
Devon Island	75°N	87°W	Ml	Koerner 77	yes	–	yes*
Ellesmere Island	80.7°N	73.1°W	Ml	Koerner & Fisher 90	yes	yes	no
Ellesmere Island	80.7°N	73.1°W	B+Is	Beltrami & Taylor 95	yes	yes	no
Gulf of Alaska	60–61°N	149°W	G+T	Calkin et al. 01	yes	yes	–
Swiss Alps (Gorner+ Grösser Aletsch Glacier)	45.8–46.5°N	7.75–8.16°E	G+Gm +Is+T	Holzhauser 97	yes	yes	no
South Georgia Island	54–55°S	36–38°W	G+Gm	Clapperton et al. 89	yes	yes	–
Southern Alps (Mueller Glacier)	43.44°S	170.06°E	G+Gm	Winkler 00	yes	–	–
Antarctica (James Ross Island)	64.22°S	57.68°W	Is	Aristarain et al. 90	yes?	–	no
Antarctica (Law Dome)	66.73°S	112.83°E	Is	Morgan 85	yes	yes	no
Antarctica (Victoria Land)	74.33°S	165.13°E	G+Gm+Is	Baroni & Orombelli 94	yes	yes	–
Antarctica (Dome C)	74.65°S	124.17°E	Is	Benoist et al. 82	yes	yes	no

Table 1 (continued)

Location	Lat.	Long.	Type	Reference	Ans. (1)	Ans. (2)	Ans. (3)
Prince William Sound, Alaska	60°N	149°W	T+G	Barclay et al. 99	yes	yes?	–
Alberta, Canada Columbia Icefield	52.2°N	117.8°W	T+In+G	Luckman et al. 97	yes	yes	yes*
N. Québec	57.73°N	76.17°W	T	Arseneault & Payette 97	–	yes	–
S. Manitoba	49.5°N	97.17°W	T	St. George & Nielsen 02	yes	–	–
Central US	33-49°N	91-109°W	T+Mp	Woodhouse & Overpeck 98	yes	yes	yes*
E. Idaho	44.1°N	114°W	T	Biondi et al. 99	yes	–	no
N. Carolina	34.5°N	78.3°W	T	Stahle et al. 88	yes	yes	no?
California (SN)	36.5-37.5°N	118.5-120.5°W	T	Graumlich 93	yes	yes	no
California (SN)	36.5-37.5°N	118.5-120.5°W	T	Scuderi 93	yes	yes	no
California (SN)	36-38°N	118-120°W	T	Swetnam 93	yes	yes	no
New Mexico	34.5°N	108°W	T	Grissino-Mayer 96	yes	yes	no
Central Eq. Pacific (NINO3.4)	5°N-5°S	160°E-150°W	T+In	Evans et al. 00	yes	yes?	no
C. Siberia (Taymir+Putoran)	72.47°N	102°E	T	Naurzbaev & Vaganov 00	yes	yes	no
Kola Peninsula	67-68°N	33-34°E	T+Is	Hiller et al 01	–	yes	–
N. Fennoscandia	68°N	22°E	T	Briffa et al. 92	yes	yes	no
NE. Italy	45°N	10°E	T	Serre-Bachet 94	yes	yes	no?
Morocco	28-36°N	2-12°W	T	Till & Guiot 90	yes	yes?	no
W. Central Asia	35-41°N	72-77°E	T	Esper et al.02b	yes	yes	no
Western Himalaya	30.5-31.2°N	78.5-80°E	T	Yadav & Singh 02	yes	–	no
Mongolia (Tarvagatay Mts.)	48.3°N	98.93°E	T	Jacoby et al. 96	yes	–	yes
Mongolia (Tarvagatay Mts.)	48.3°N	98.93°E	T+D	D'Arrigo et al. 01	yes	yes	yes
N. Patagonia (Rio Alerce, Argentina)	41.17°S	71.77°W	T	Villalba 90	yes	yes	no
S. Chile (Lenca)	41.55°S	72.6°W	T	Lara & Villalba 93	yes	no	no
S. South America	33-55°S	60-75°W	T+G	Villalba 94	yes	yes	no
W. Tasmania	42°S	146.5°E	T	Cook et al. 00	no	yes	yes?
New Zealand	35-48°S	167-177°E	T	D'Arrigo et al. 98	yes	–	no
N. Scandinavia	68°N	20°E	T+G	Karlén 98	yes	yes	no
California (SN)	38°N	110°W	Ts	Stine 94	–	yes	no
California (SN)	37.5°N	119.45°W	Ts	Stine 94	–	yes	no
California (SN)	38.38°N	119.45°W	Ts	Stine 94	–	yes	no
California (SN)	38.85°N	120.47°W	Ts	Stine 94	–	yes	no
Patagonia	48.95°S	71.43°W	Ts	Stine 94	–	yes	–
Patagonia	50.47°S	72.97°W	Ts	Stine 94	–	yes	–

Table 1 (continued)

Location	Lat.	Long.	Type	Reference	Ans. (1)	Ans. (2)	Ans. (3)
NW Michigan (L. Marion)	45°N	85°W	Po	Bernabo 81	yes	yes	yes?
Qinghai-Tibetan P. (Dunde Ice Cap)	38.1°N	96.4°E	Po	Liu et al. 98	yes	yes	yes*?†
NE China (Maili)	42.87°N	122.87°E	Po	Ren 98	–	yes	–
NE China (Hangzhou)	30-33°N	105-122°E	Pf	Zhang 94	–	yes	–
China (Taibai Mt.)	33.97°N	107.73°E	Pf+Po	Tong et al. 96	yes	yes	no
Himalaya	28.38°N	85.72°E	Is	Thompson et al. 00	yes	no	yes
Himalaya (Dasuopu Glacier)	28.38°N	85.72°E	Ic	Thompson et al. 00	yes	–	yes
Guliya Ice Cap	35.2°N	81.5°E	Ic+Is	Thompson et al. 95	yes	yes	no
E. China	30-40°N	100-120°E	Ic+D	Shi et al. 99	yes	yes	yes?
W. China (Guliya Cap)	35.2°N	81.5°E	Ic+D	Shi et al. 99	yes	yes	yes?
Quelccaya Ice Cap	13.93°S	70.83°W	Is+Ic	Thompson et al. 86	yes	yes?	no
Antarctica (Siple Station)	75.92°S	84.25°W	Ic+Is	Mosley-Thompson 95	no	–	no
Antarctica (Dyer Plateau)	70.67°S	64.88°W	Ic+Is	Thompson et al. 94	yes?	–	yes
Antarctica (Dronning Maud Land)	76°S	8.05°W	Ic+Is	Karlöf et al. 00	yes	yes?	no?
South Pole	90°S	–	Ic	Mosley-Thompson & Thompson 82	yes	yes?	no

† For the Dunde ice cap, Thompson et al. (1989) noted that, according to the $\delta^{18}O$ climate proxy, the decades of the 1940s, 1950s and 1980s are at least as warm as the Holocene maximum of 6000 to 8000 years ago. In order to confirm Thompson et al. (1989)'s original statement, please consider Figure 6 of Thompson (2000), because the claim that 1930s–1980s is the warmest of the last 6000–8000 years ago is not clear from any figure in Thompson et al. (1989). But the main warming of the 1940s–1950s occurred before significant rise of anthropogenic CO_2 in the air. This is why we added a question mark plus an asterisk to this entry.

Table 1 (continued)

Location	Lat.	Long.	Type	Reference	Ans. (1)	Ans. (2)	Ans. (3)
N. Atlantic	54.27°N	16.78°W	Sd	Bond et al. 97	yes	yes	no?
N. Atlantic	44.5°N	46.33°W	Sd	Bond et al. 99	yes	yes	no
N. Atlantic	56.37°N	27.81°W	Sd	Bianchi & McCave 99	yes	yes	no?
N. Ellesmere Island	81°N	80°W	Sd+Lf	Lamoureux & Bradley 96	yes	yes	–
SW. Baltic Sea (Bornholm Basin)	55.38°N	15.4°E	Sd+Is	Andrén et al. 00	yes	yes	no
S. China (Huguangyan L.)	21.15°N	110.28°E	Lf+Is	Chu et al. 02	yes	yes	–
NW. Finland (L. Toskaljavri)	69.2°N	21.47°E	Lf+Po+T	Seppä & Birks 02	yes	yes	–
N. Fennoscandia (L. Tsuolbmajavri)	68.68°N	22.08°E	Lf	Korhola et al. 00	yes	yes	no
S. Finland (L. Laihalampi)	61.49°N	26.08°E	Lf+Po	Heikkilä & Seppä 03	–	no?	no
Switzerland (L. Neuchatel)	47°N	6.55°W	Lf+Is	Filippi et al. 99	yes	yes	yes?
W. Ireland	53.53°N	9.93°W	Is	Blackford & Chambers 95	yes	yes	–
Bermuda Rise	32.17°N	64.5°W	Is	Keigwin 96	yes	yes	no
Chesapeake Bay	37-38.4°N	76.1°W	Sd	Verardo et al. 98	yes	yes	–
Chesapeake Bay	38-38.9°N	76.22-76.4°W	Sd+Is	Cronin et al. 03	yes	yes	yes?
NW Alaska (Farewell L.)	62.55°N	153.63°W	Lf+Is	Hu et al. 01	yes	yes	no
W Canada (Pine L.)	52.07°N	113.45°W	Lf	Campbell et al. 98	yes	yes	no
S. Dakota (Pickerel L.)	45.51°N	97.27°W	Lf	Dean & Schwalb 00	yes	yes	no
N. Dakota (Moon L.)	46.85°N	98.16°W	Lf	Laird et al. 96	yes	yes	no
N. Dakota (Rice L.)	48.01°N	101.53°W	Lf	Yu & Ito 99	yes	yes	no
Yellowstone P. (Lamar Cave)	44.56°N	110.24°W	Pf+Is	Hadly 96	yes	yes	no
Colorado Plateau (L. Canyon)	37.42°N	110.67°W	Lf+Gm+Is	Pederson 00	yes	yes	–
NE Colorado	40-41°N	102-105°W	Gm+Is+D	Madole 94	–	yes	no
SW US (Colorado +Arizona)	34-37.5N –	105-112W	Lf+Is	Davis 94	–	yes	no?
SW US	32-39°N	109-114°W	Lf+Gm	Ely et al. 93	yes	yes	no
California (White Mts.)	37.43°N	118.17°W	Is	Feng & Epstein 94	yes	yes	–
California (L. Owen)	36°N	118.17°W	Is	Li et al. 00	yes	yes	no
Yucatan Peninsula (L. Chichancanab)	20°N –	88.4°W	Lf+Is	Hodell et al. 01	yes	yes	–
Cariaco Basin	11°N	65°W	Sd	Black et al. 99	yes	yes	no
Cariaco Basin	10.71°N	65.17°W	Sd+Is	Haug et al. 01	yes	yes	–
S. Florida	24.95°N	80.55°W	Is	Druffel 82	yes	–	–
NE. Carribean	17.89°N	66.60°W	Sd+Is	Nyberg et al. 02	yes	yes	no
SW. Puerto Rico	18.12°N	67.09°W	Is	Winter et al. 00	yes	–	–

Table 1 (continued)

Location	Lat.	Long.	Type	Reference	Ans. (1)	Ans. (2)	Ans. (3)
NW Scotland (Assynt)	58.11°N	5.06°W	Sp	Proctor et al. 00	yes	yes	no
SW. Ireland	52.5°N	9.25°W	Sp	McDermott et al. 01	yes	yes	–
NW. Germany (Sauerland Mtn)	51.43°N	7.78°W	Sp+Is	Niggemann et al. 03	yes	yes	–
NE China (Jinchuan)	42.3°N	126.37°E	Is	Hong et al. 00	yes	yes	no
NE China (Shihua Cave)	39.8°N	115.9°E	Sp+Is	Li et al. 98, Ku & Li 98	yes	yes	no
C. China (Buddha Cave)	33.67°N	109.08°E	Sp+Is	Paulsen et al. 03	yes	yes	no?
S. Japan (Yakushima Is.)	30.33°N	130.5°E	Is	Kitagawa & Matsumoto 95	yes	yes	no
N. India (Pahalgam)	34.02°N	75.20°E	Is	Ramesh 93	yes	–	–
S. India (Nilgiris)	10-10.5°N	77°E	Is	Ramesh 93	–	yes	–
E. Africa (L. Malawi)	10°S	35°E	Lf	Johnson et al. 01	yes	–	no
E. Africa (L. Naivasha)	0.46°S	36.21°E	Lf	Verschuren et al. 00	yes	yes	no
W. Africa (Cap Blanc)	20.75°N	18.58°W	Is	deMenocal et al. 00	yes	yes	no
S. Africa	19-35°S	10-33°E	Mp	Tyson & Lindesay 92	yes	yes	no
S. Africa (Nelson Bay Cave)	34°S	23°E	Is	Cohen & Tyson 95	yes	–	no
S. Africa (Makapansgat)	24.54°S	29.25°E	Sp	Tyson et al. 00	yes	yes	no
S. Oman	16.93°N	54.00°E	Sp+Is	Burns et al. 02	yes	–	no
W. Pacific	6.3°N	125.83°E	Sd+Is	Stott 02, Stott et al. 02	yes	yes	no
N. New Zealand (Waitomo)	38.27°S	175°E	Sp	Williams et al. 99	yes	–	–
S. New Zealand (Nelson)	40.67°S	172.43°E	Sp	Wilson et al. 79	yes	yes	no
S. America (multiple regions)	33-38°S	59.3-67°W	Mp	Iriondo 99	yes	yes	–
C. Argentina	29.5-35°S	61.75-65.75°W	Gm+D	Carignano 99	yes	yes	no
C. Argentina	28-36°S	61-67°W	G+Mp	Cioccale 99	yes	yes	no
NW. Argentina	26.5°S	68.09°W	Sd+Is	Valero-Garcés et al. 00	yes	–	–
W. Antarctica (E. Bransfield Basin)	61.98°S	55.95°W	Sd+Is	Khim et al. 02	yes	yes	–
W. Antarctica (Palmer Deep)	64.86°S	64.21°W	Sd	Domack et al. 01	yes	yes	no
W. Antarctica (Siple Dome)	81.65°S	148.81°W	Is	Kreutz et al. 97	yes	–	no

yes*: the warming or extreme excursion peaked around 1920s–1950s before any significant anthropogenic CO_2 release to air anwers ending with a question mark (?) refers to indecision

B: Borehole

Cl: Cultural

D: Documentary

G: Glacier advance or retreat

Gm: Geomorphology

In: Instrumental

Is: Isotopic analysis from lake sedimentary or ice cores, tree or peat celluloses, corals, stalagmite or biological fossils

Ic: Net ice accumulation rate, including dust or chemical counts

Lf: Lake fossils and sediments; river sediments

Ml: Melt layers in ice cores

Mp: Multiproxy and can be any combination of the proxies list here

Pf: Phenological and Paleontological fossils

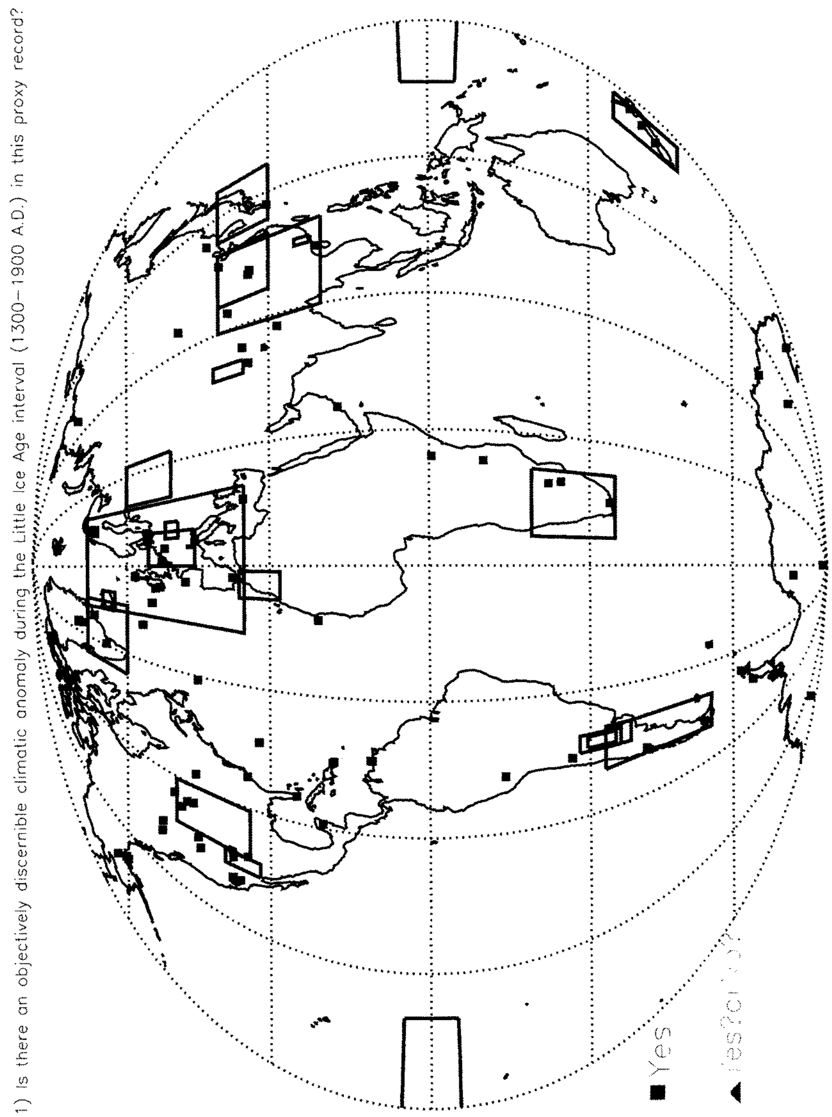
Po: Pollen

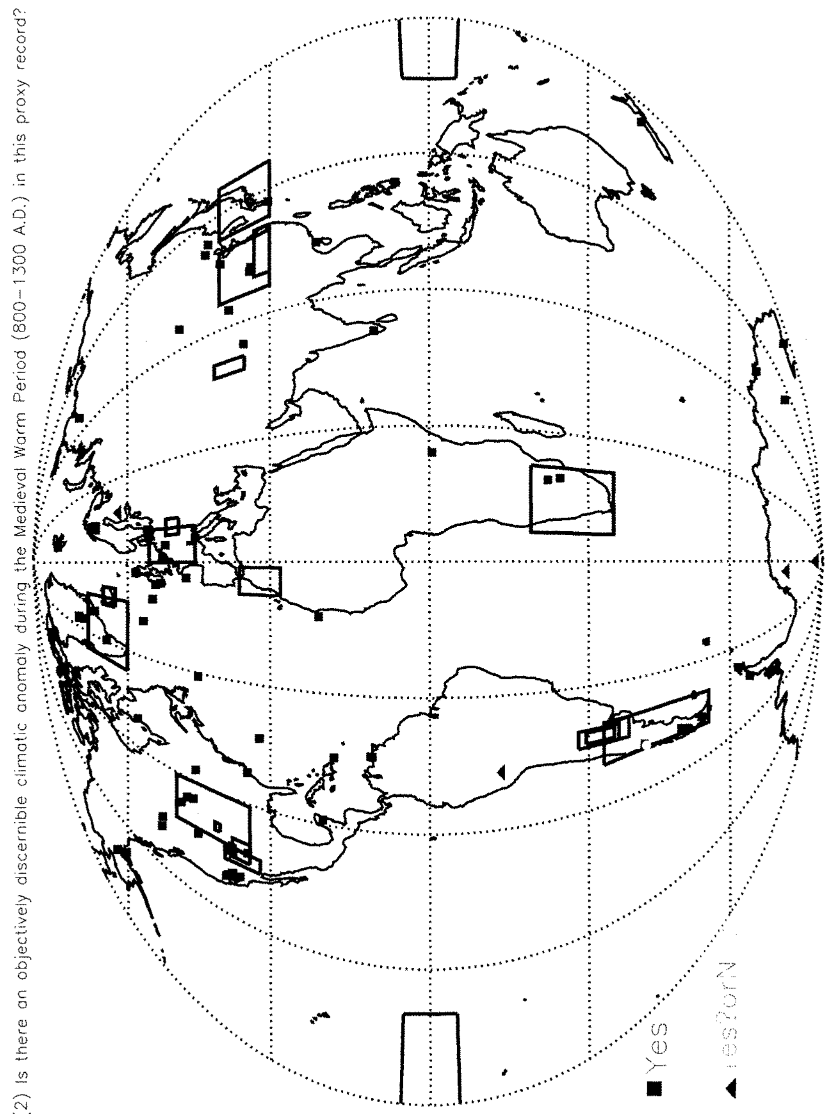
Sd: Seafloor sediments

Sp: Speleothem isotopic or luminescent analysis

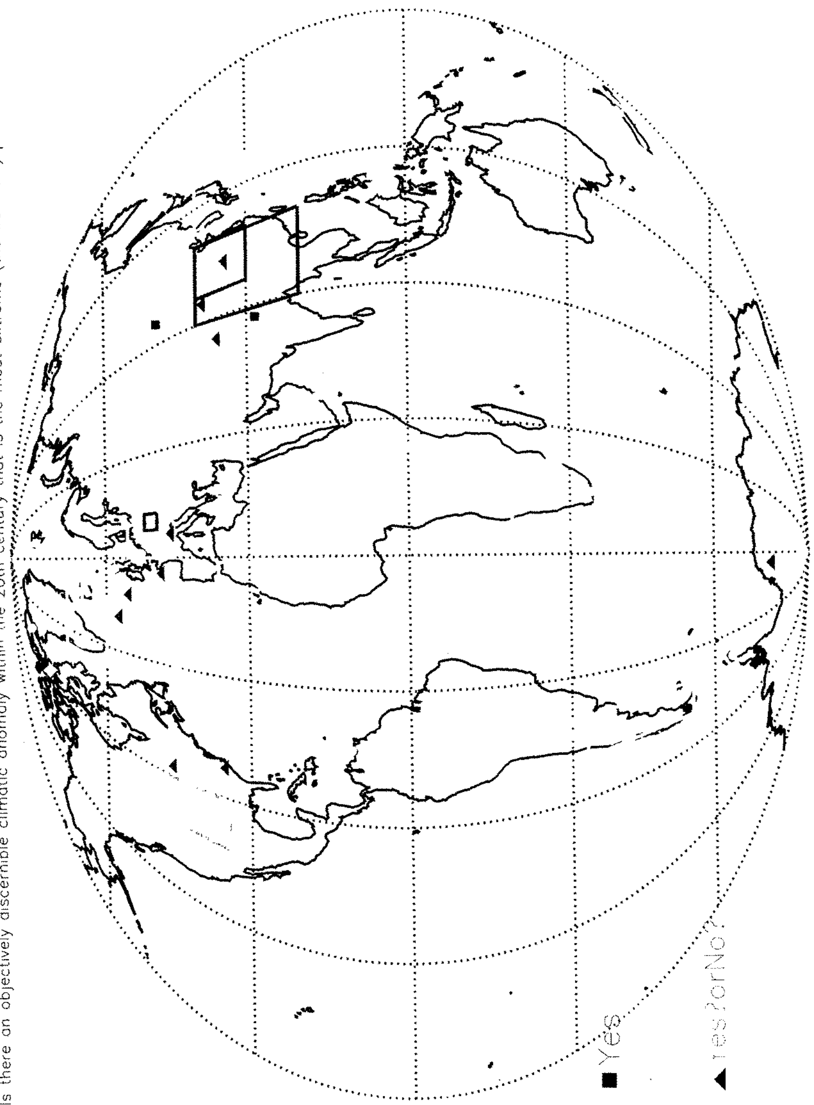
T: Tree ring growth, either ring width or maximum latewood density, including shifting tree line positions

Ts: Tree stumps in lakes, marshes and streams


(1) Is there an objectively discernible climatic anomaly during the Little Ice Age interval (1300–1900 A.D.) in this proxy record?
Yes


(2) Is there an objectively discernible climatic anomaly during the Medieval Warm Period (800–1300 A.D.) in this proxy record?
Yes

(3) Is there an objectively discernible climatic anomaly within the 20th century that is the most extreme (the warmest) period in the record?



Vol. 23: 89–110, 2003 | CLIMATE RESEARCH Clim Res | Published January 31

Proxy climatic and environmental changes of the past 1000 years

Willie Soon[1,2,*], Sallie Baliunas[1,2]

[1]Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, MS 16, Cambridge, Massachusetts 02138, USA
[2]Mount Wilson Observatory, Mount Wilson, California 91023, USA

ABSTRACT: The 1000 yr climatic and environmental history of the Earth contained in various proxy records is reviewed. As indicators, the proxies duly represent local climate. Because each is of a different nature, the results from the proxy indicators cannot be combined into a hemispheric or global quantitative composite. However, considered as an ensemble of individual expert opinions, the assemblage of local representations of climate establishes both the Little Ice Age and Medieval Warm Period as climatic anomalies with worldwide imprints, extending earlier results by Bryson et al. (1963), Lamb (1965), and numerous intervening research efforts. Furthermore, the individual proxies can be used to address the question of whether the 20th century is the warmest of the 2nd millennium locally. Across the world, many records reveal that the 20th century is probably not the warmest nor a uniquely extreme climatic period of the last millennium.

KEY WORDS: Paleoclimate proxies · Climate change · Environmental change · Little Ice Age · Medieval Warm Period

1. INTRODUCTION

Are the Little Ice Age and Medieval Warm Period widespread climatic anomalies? Lamb (1965) wrote, '[M]ultifarious evidence of a meteorological nature from historical records, as well as archaeological, botanical and glaciological evidence in various parts of the world from the Arctic to New Zealand . . . has been found to suggest a warmer epoch lasting several centuries between about A.D. 900 or 1000 and about 1200 or 1300. . . . Both the "Little Optimum" in the early Middle Ages and the cold epochs [i.e. "Little Ice Age"], now known to have reached its culminating stages between 1550 and 1700, can today be substantiated by enough data to repay meteorological investigation. . . . It is high time therefore to marshal the climatic evidence and attempt a quantitative evidence' (p. 14–15). Research on large-scale patterns of climate change continued with vigor.

Jones et al. (1998) tentatively concluded that while a Little Ice Age cooling existed, little evidence could be found to support or reject a medieval warming. But the updated composite tree-ring summer temperature curve in Fig. 1 of Briffa (2000) shows evidence for an anomalously warm interval from about 950 to 1100 in the northern high-latitude zone, which coincides with Lamb's Medieval Warm Period. Also, a similar early warm period appears prominently in the averaged tree ring chronologies carefully selected and processed from 14 sites spreading over 30 to 70°N (Esper et al. 2002).

Those results are but a few of many that have become available since Lamb's analysis. Given advancements in retrieval of information from and extension of surface coverage for the proxies, we review the accumulated evidence on regional climatic anomalies over the last 1000 yr. We also recommend Ogilvie & Jónsson (2001), who recently provided the most authoritative discussion on the historical development of the long-standing debates on the climatic nature of the Medieval Warm Period and Little Ice Age, especially concerning the North Atlantic, including Iceland.

*Email: wsoon@cfa.harvard.edu

2. WORKING DEFINITIONS

What are the regional patterns of climatic change over the last 1000 yr? Accurate results could serve as benchmarks for 20th century global average warming seen in the surface thermometer records and as physical constraints for theories or mechanisms of climate change on timescales of decades to centuries.

The proxies used to study climatic change over the last 1000 yr are addressed individually and therefore locally because they differ in nature too greatly to be quantitatively averaged or compared. To make progress, we consider 3 questions of many individual climate proxies:

(1) Is there an objectively discernible climatic anomaly during the Little Ice Age interval, defined as 1300–1900? This broad period in our definition derives from historical sea-ice, glaciological and geomorphological studies synthesized in Grove (2001a,b) and Ogilvie & Jónsson (2001).

(2) Is there an objectively discernible climatic anomaly during the Medieval Warm Period, defined as 800–1300? This definition is motivated by e.g. Pfister et al. (1998) and Broecker (2001), and is a slight modification of Lamb's original study (1965).

(3) Is there an objectively discernible climatic anomaly within the 20th century that is the most extreme (the warmest, if such information is retrievable) period in the record? An important consideration in answering this question is to distinguish the cases for which the 20th century warming began early in the century versus after the 1970s, as recorded by surface thermometers. This is necessary if temperature changes are to be related to anthropogenic forcing inputs like increased atmospheric carbon dioxide content.

Anomaly is simply defined as a period of more than 50 yr of sustained warmth, wetness or dryness, within the stipulated interval of the Medieval Warm Period, or a 50 yr or longer period of cold, dryness or wetness within the stipulated Little Ice Age. We define anomaly in the 20th century within each proxy in the same way. The surface instrumental record of the 20th century contains 3 distinct, multidecadal trends: early-century warming, mid-century cooling and late-century warming. But that knowledge comes from instrumental thermometry with its high time resolution and other biases, which preclude a direct comparison to the proxies (which have their own biases). Our goal here is to compare the 20th century objectively with more extended past changes than is available from thermometry. Given the biases of each proxy, Question (3) was answered by asking if, within each proxy record, there was an earlier (pre-20th century) 50 yr interval warmer (or more extreme, in the case of precipitation) than any 50 yr average within the 20th century.

Question (3) differs from Questions (1) and (2). Question (3) looks for a 50 yr anomaly within the 20th century compared to any other anomaly throughout the period of a proxy record. Questions (1) and (2) look for a 50 yr anomaly within the previously suggested 500 and 600 yr intervals of the Medieval Warm Period and Little Ice Age, respectively. Note that in the case of Question (3), we treat the definition of a 50 yr or more period of sustained anomaly in the 20th century no different from that of any prior century. Thus, if a sustained warm anomaly were identified and happened to reside in the Medieval Warm Period and appeared warmer than an anomaly found in the 20th century, then we would assign 'No' to Question (3). Similarly, a proxy record may show both that the 20th century anomaly is the most extreme (warmest) and that the Medieval Warm Period exists. In answering Question (3), the existence of the Medieval Warm Period or Little Ice Age is not considered, because they are assessed independently, in Questions (1) and (2).

We started with the framework of past researchers, namely, the suggested existence of the Medieval Warm Period and Little Ice Age. The goal of the study is to deduce the geographical nature of climatic and environmental conditions during these periods. Distinguishing the 20th century as a separate period is largely a practical bias because of the interest in the role of human activity on Earth's climate and environment.

Another important consideration is that temporary regional cooling (or shift between wet and dry condition) may occur on decadal, but not multidecadal, timescales during the Medieval Warm Period, and similarly, occasional, short-lived regional warming (or shift between wet and dry condition) may occur during the Little Ice Age, as indicated by Grove (2001a,b). Thus the terms Medieval Warm Period and Little Ice Age should indicate persistent but not necessarily constant warming or cooling, respectively, over broad areas (see also Stine 1998, Luckman 2000, Grove 2001a,b, Ogilvie & Jónsson 2001, Esper et al. 2002). Stine (1998) suggests that more appropriate terms may be 'Little Ice-Age Climatic Anomaly' and 'Medieval Climatic Anomaly'. It is also noteworthy that the definitions of discernible, persistent climate anomalies for the Little Ice Age and Medieval Warm Period may include not only distinct changes in the mean but also changes in the multidecadal variance (Ogilvie & Jónsson 2001). Through the microscope of daily and regional spatial scale variability, it is important to recognize that even the relation between multidecadal mean temperature and its daily variability may undergo significant non-stationary changes (e.g. Knappenberger et al. 2001, who document those time-dependent changes in temperature variability across the United States).

Also, from a combination of field evidence and modeling based on the understanding from synoptic climatology, Bryson & Bryson (1997) demonstrated how local and regional factors (for horizontal spatial distances as small as 100 km) can produce significantly different precipitation histories for 2 Near East stations (Jerusalem and Kameshli, Syria) and for 2 stations in the Cascade Range of Oregon (mountainous versus coastal-like microclimate sites).

Our classification of a widespread anomaly based on multidecadal persistence at many locales rests on good precedent. For example, the modern globally averaged surface warming inferred from thermometer readings includes large-scale cooling trends over both the Greenland/Labrador Sea area and the eastern region of the United States (30 to 45° N; 80 to 110° W; see Hansen et al. 1999, Robinson et al. 2001) or the Antarctic continent (e.g. Doran et al. 2002) in the last 30 to 50 yr. Another example is relative warmth during the Little Ice Age and relative cool during the Medieval Warm Period seen in the borehole record of reconstructed temperature at Taylor Dome, Antarctica (77.8° S, 158.72° E, elevation 2374 m), compared to results from Greenland's borehole (see Clow & Waddington 1999), which do not show those features. These variations are of short duration compared to the anomaly, and of limited regional extent.

3. APPROACH

Table 1 and Figs. 1 to 3 summarize the answers to the questions posed here about local climatic anomalies. For Questions (1) and (2), we answered 'Yes' if the proxy record showed a period longer than 50 yr of cooling, wetness or dryness during the Little Ice Age, and similarly for a period of 50 yr or longer for warming, wetness or dryness during the Medieval Warm Period. A dash indicates that either there is no expert opinion, or that the proxy record does not cover the period in question. A 'Yes?' or 'No?' answer means that the original expert opinion was Yes or No, but that it does not match our criteria; for example, if the interval of warmth during the Medieval Warm Period were too short by our definition to be 'Yes', it would merit 'Yes?'. In several cases in the 20th century, a 'Yes[a]' designation was assigned for the answer to Question (3) in order to highlight the fact that the 20th century warming first occurred early in the century, ca. 1920–1950, when the air's content of anthropogenic CO_2 was still cumulatively small.

A global association for the Little Ice Age or Medieval Warm Period is premature because proxy data are geographically sparse and either or both phenomena could be multi-phased events acting under distinct local and regional constraints and modes. Bradley & Jones (1993) and Hughes & Diaz (1994) initiated and strongly held the view that the phenomena were not global, but Grove (1996; see especially p. 51 to 54) disagrees. However, in the traditionally data-rich areas of Western Europe or the Northern Atlantic including Iceland and Greenland, both the Little Ice Age and Medieval Warm Period do exist as distinct climate anomalies (Pfister et al. 1998, Grove 2001a, Ogilvie & Jónsson 2001). No objective proof discredits the existence of those phenomena in other regions. Thus, like other researchers (Lamb 1965, Porter 1986, Grove 1996, Kreutz et al. 1997), we start with previously indicated periods of Little Ice Age and Medieval Warm Period, and ask whether they are widespread, teleconnected events that need not necessarily last throughout the defined periods. The terms Medieval Warm Period and Little Ice Age remain practical and viable, especially considering the potential extension of the concept to past and future climatic events that are 'similar or equivalent' in physical scope (e.g. Bond et al. 1997, 1999).

Current knowledge on the diverse range of local climatic behavior suggests that the Medieval Warm Period and the Little Ice Age are not expected to be homogeneous and sustained. To define the beginning and end dates of these climate anomalies requires better understanding (for the Little Ice Age see Porter 1981, 1986, Kaser 1999, Grove 2001a,b, Luckman 2000, Schuster et al. 2000, Winkler 2000, Ogilvie & Jónsson 2001, Hendy et al. 2002). Imprecise data on the beginning and end of both events contributes in part to confusion about the phenomena. For example, Ogilvie & Farmer (1997) have commented that Lamb's suggestion of a Medieval Warm Period may not be supported by documentary data even for England because their extensive studies based on a historical dataset showed that England suffered relatively cold winters from 1260 to 1360. However, as that period is near the transition between the Medieval Warm Period and Little Ice Age defined in this study, this fact does not strongly contradict our results. Evidence based primarily on glacier activity points to both a poorly defined beginning and end of the Medieval Warm Period, while the Little Ice Age interval seems to have had a gradual beginning but more abrupt end. Although the notion of the Medieval Warm Period or Little Ice Age with sharply defined transitions may be a convenient one, it is probably a non-physical construct because of large regional differences in the timing of both phenomena. As suggested by Grove (2001a), an inhomogeneous climate pattern (though not necessarily an analog) can already be identified in the 20th century warm intervals as defined by instrumental records.

Table 1. A full list of paleoclimatic proxies that have sufficient length of continuous records to entertain the 3 specifc questions: (1) Is there an objectively discernible climatic anomaly during the Little Ice Age interval (A.D. 1300–1900) in this proxy record? (2) Is there an objectively discernible climatic anomaly during the Medieval Warm Period (A.D. 800–1300) in this proxy record? (3) Is there an objectively discernible climatic anomaly within the 20th century that is the most extreme (the warmest, if such information is available) period in the record? A question mark after an answer indicates uncertainty or indecision. Symbols for the type of climate proxy used include B: borehole; Cl: cultural; D: documentary; G: glacier advance or retreat; Gm: geomorphology; In: instrumental; Is: isotopic analysis from lake sedimentary or ice cores, tree or peat celluloses, corals, stalagmite or biological fossils; Ic: net ice accumulation rate, including dust or chemical counts; Lf: lake fossils and sediments; river sediments; Ml: melt layers in ice cores; Mp: multiproxy (any combination of proxies listed here); Pf: phenological and paleontological fossils; Po: pollen; Sd: seafloor sediments; Sp: speleothem isotopic or luminescent analysis; T: tree ring growth, either ring width or maximum latewood density, including shifting tree line positions; Ts: tree stumps in lakes, marshes and streams

Location	Latitude	Longitude	Type	Source	Answer		
					(1)	(2)	(3)
Worldwide	–	–	Mp	Mann et al. (1999)	Yes	No	Yes[a]
Arctic wide	–	–	Mp	Overpeck et al. (1997)	Yes	–	Yes[a]
Worldwide	–	–	Mp	Crowley & Lowery (2000)	Yes	No	Yes[a]
Worldwide	–	–	Mp	Jones et al. (1998)	Yes	No	Yes[a]
Worldwide	–	–	T	Briffa (2000)	Yes	No	Yes[a]
Worldwide	–	–	T	Briffa et al. (2001)	Yes	–	Yes[a]
Worldwide	–	–	T	Jones et al. (2001)	Yes	–	Yes[a]
NH mid-latitude	–	–	T	Esper et al. (2002)	Yes	Yes	No
Worldwide	–	–	Mp	Lamb (1977, 1982)	Yes	Yes	–
Worldwide	–	–	G + Is	Porter (1986)	Yes	Yes	–
Worldwide	–	–	G	Grove & Switsur (1994)	–	Yes	–
Worldwide	–	–	T+G+D	Hughes & Diaz (1994)	Yes	No?[b]	No?[b]
Worldwide	–	–	Mp	Grove (1996)	–	Yes	–
Worldwide	–	–	B	Huang et al. (1997)	Yes	Yes	No
Worldwide	–	–	D	Perry & Hsu (2000[c])	Yes	Yes	No
Worldwide	–	–	D	deMenocal (2001)	Yes	Yes	–
Americas	–	–	Ts+Gm+Mp	Stine (1998)	–	Yes	–
N. Atlantic (Iceland)	63–66° N	14–24° W	Mp	Ogilvie et al. (2000)			
				Ogilvie & Jónsson (2001)	Yes	Yes	No
N. Atlantic (S. Greenland)	60–70° N	20–55° W	Mp	Ogilvie et al. (2000)	Yes	Yes	No
W. Europe	45–54° N	0–15° E	Mp	Pfister et al. (1998)	Yes	Yes	No
N. Atlantic (Europe)	35–70° N	25° W–30° E	In+D	Luterbacher et al. (2000)	Yes	–	–
Central England	52° N	2° E	In+D	Lamb 65, Manley (1974)	Yes	Yes	No
S. Spain	37.30° N	4.30° W	In+D	Rodrigo et al. (2000)	Yes	–	No
Crete Is.	35.15° N	25.00° E	D	Grove & Conterio (1995)	Yes	–	No
Mid-Russia	50–60° N	30–50° E	In+D	Borisenkov (1995)	Yes	–	–
Czech Republic	48.5–51.2° N	12–19° E	In+B	Bodri & Cermák (1999)	Yes	Yes	Yes?
S. USA	37–38° N	107.5–109.5° W	Pf+Cl+D	Petersen (1994)	Yes	Yes	–
E. China (Guang Dong Prov.)	22–25° N	112–114.3° E	D	Chan & Shi (2000)	Yes	–	–
E. China-wide	20–40° N	90–120° E	D	Song (2000)	Yes	–	No?
Japan	30–40° N	125–145° E	D	Tagami (1993, 1996)	Yes	Yes	No
S. Africa	22.2° S	28.38° E	Cl	Huffman (1996)	Yes	Yes	–
E. Greenland (Nansen Fjord)	68.3° N	29.7° W	Is	Jennings & Weiner (1996)	Yes	Yes	No
C. Greenland (Crête)	71.12° N	37.32° W	Is	Dansgaard et al. (1975)	Yes	Yes	No
C. Greenland (GRIP)	72.6° N	37.6° W	B	Dahl-Jensen et al. (1998)	Yes	Yes	No
S. Greenland (Dye 3)	65.2° N	43.8° W	B	Dahl-Jensen et al. (1998)	Yes	Yes	No
C. Greenland (GISP2)	72.58° N	38.5° W	Ic+Ml	Meese et al. (1994)	Yes	Yes	No
C. Greenland (GISP2)	72.58° N	38.5° W	Is	Stuiver et al. (1995)	Yes	Yes	No
Svalsbard	79° N	15° E	Ml	Tarussov (1995)	Yes	Yes	No
Devon Island	75° N	87° W	Ml	Koerner (1977)	Yes	–	Yes[a]
Ellesmere Island	80.7° N	73.1° W	Ml	Koerner & Fisher (1990)	Yes	Yes	No
Ellesmere Island	80.7° N	73.1° W	B+Is	Beltrami & Taylor (1995)	Yes	Yes	No
Gulf of Alaska	60–61° N	149° W	G+T	Calkin et al. (2001)	Yes	Yes	–
Swiss Alps (Gorner Glacier)	45.8–46.5° N	7.75–8.16° E	G+Gm	Holzhauser (1997)	Yes	Yes	No
(Grosser Aletsch Glacier)	45.8–46.5° N	7.75–8.16° E	Is+T	Holzhauser (1997)	Yes	Yes	No
South Georgia Island	54–55° S	36–38° W	G+Gm	Clapperton et al. (1989)	Yes	Yes	–
Southern Alps (Mueller Glacier)	43.44° S	170.06° E	G+Gm	Winkler (2000)	Yes	–	–
Antarctica (James Ross Island)	64.22° S	57.68° W	Is	Aristarain et al. (1990)	Yes?	–	No
Antarctica (Law Dome)	66.73° S	112.83° E	Is	Morgan (1985)	Yes	Yes	No
Antarctica (Victoria Land)	74.33° S	165.13° E	G+Gm+Is	Baroni & Orombelli (1994)	Yes	Yes	–

Table 1 (continued)

Location	Latitude	Longitude	Type	Source	Answer (1)	Answer (2)	Answer (3)
Antarctica (Dome C)	74.65° S	124.17° E	Is	Benoist et al. (1982)	Yes	Yes	No
Prince William Sound, Alaska	60° N	149° W	T+G	Barclay et al. (1999)	Yes	Yes?	–
Alberta, Canada Columbia Icefield	52.2° N	117.8° W	T+In+G	Luckman et al. (1997)	Yes	Yes	Yes[a]
N. Québec	57.73° N	76.17° W	T	Arseneault & Payette (1997)	–	Yes	–
Central US	33–49° N	91–109° W	T+Mp	Woodhouse & Overpeck (1998)	Yes	Yes	Yes[a]
E. Idaho	44.1° N	114° W	T	Biondi et al. (1999)	Yes	–	No
N. Carolina	34.5° N	78.3° W	T	Stahle et al. (1988)	Yes	Yes	No?
California (SN)	36.5–37.5° N	118.5–120.5° W	T	Graumlich (1993)	Yes	Yes	No
California (SN)	36.5–37.5° N	118.5–120.5° W	T	Scuderi (1993)	Yes	Yes	No
California (SN)	36–38° N	118–120° W	T	Swetnam (1993)	Yes	Yes	No
New Mexico	34.5° N	108° W	T	Grissino-Mayer (1996)	Yes	Yes	No
Central Eq. Pacific (NINO 3.4)	5° N–5° S	160° E–150° W	T+In	Evans et al. (2000)	Yes	Yes?	No
C. Siberia (Taymir + Putoran)	72.47° N	102° E	T	Naurzbaev & Vaganov (2000)	Yes	Yes	No
Kola Peninsula	67–68° N	33–34° E	T+Is	Hiller et al. (2001)	–	Yes	–
N. Fennoscandia	68° N	22° E	T	Briffa et al. (1992)	Yes	Yes	No
NE. Italy	45° N	10° E	T	Serre-Bachet (1994)	Yes	Yes	No?
Morocco	28–36° N	2–12° W	T	Till & Guiot (1990)	Yes	Yes?	No
Mongolia (Tarvagatay Mts.)	48.3° N	98.93° E	T	Jacoby et al. (1996)	Yes	–	Yes
Mongolia (Tarvagatay Mts.)	48.3° N	98.93° E	T+D	D'Arrigo et al. (2001)	Yes	Yes	Yes
N. Patagonia (Rio Alerce, Argentina)	41.17° S	71.77° W	T	Villalba (1990)	Yes	Yes	No
S. Chile (Lenca)	41.55° S	72.6° W	T	Lara & Villalba (1993)	Yes	No	No
S. South America	33–55° S	60–75° W	T+G	Villalba (1994)	Yes	Yes	No
W. Tasmania	42° S	146.5° E	T	Cook et al. (2000)	No	Yes	Yes?
New Zealand	35–48° S	167–177° E	T	D'Arrigo et al. (1998)	Yes	–	No
N. Scandinavia	68° N	20° E	T+G	Karlén (1998)	Yes	Yes	No
California (SN)	38° N	110° W	Ts	Stine (1994)	–	Yes	No
California (SN)	37.5° N	119.45° W	Ts	Stine (1994)	–	Yes	No
California (SN)	38.38° N	119.45° W	Ts	Stine (1994)	–	Yes	No
California (SN)	38.85° N	120.47° W	Ts	Stine (1994)	–	Yes	No
Patagonia	48.95° S	71.43° W	Ts	Stine (1994)	–	Yes	–
Patagonia	50.47° S	72.97° W	Ts	Stine (1994)	–	Yes	–
NW Michigan (L. Marion)	45° N	85° W	Po	Bernabo (1981)	Yes	Yes	Yes?
Qinghai-Tibetan Plat. (Dunde Ice Cap)	38.1° N	96.4° E	Po	Liu et al. (1998)	Yes	Yes	Yes?[a,d]
NE China (Maili)	42.87° N	122.87° E	Po	Ren (1998)	–	Yes	–
NE China (Hangzhou)	30–33° N	105–122° E	Pf	Zhang (1994)	–	Yes	–
China (Taibai Mt.)	33.97° N	107.73° E	Pf+Po	Tong et al. (1996)	Yes	Yes	No
Himalaya	28.38° N	85.72° E	Is	Thompson et al. (2000)	Yes	No	Yes
Himalaya (Dasuopu Glacier)	28.38° N	85.72° E	Ic	Thompson et al. (2000)	Yes	–	Yes
Guliya Ice Cap	35.2° N	81.5° E	Ic+Is	Thompson et al. (1995)	Yes	Yes	No
E. China	30–40° N	100–120° E	Ic+D	Shi et al. (1999)	Yes	Yes	Yes?
W. China (Guliya Cap)	35.2° N	81.5° E	Ic+D	Shi et al. (1999)	Yes	Yes	Yes?
Quelccaya Ice Cap	13.93° S	70.83° W	Is+Ic	Thompson et al. (1986)	Yes	Yes?	No
Antarctica (Siple Station)	75.92° S	84.25° W	Ic+Is	Mosley-Thompson (1995)	No	–	No
Antarctica (Dyer Plateau)	70.67° S	64.88° W	Ic+Is	Thompson et al. (1994)	Yes?	–	Yes
Antarctica (Dronning Maud Land)	76° S	8.05° W	Ic+Is	Karlöf et al. (2000)	Yes	Yes?	No?
South Pole	90° S	–	Ic	Mosley-Thompson & Thompson (1982)	Yes	Yes?	No
N. Atlantic	54.27° N	16.78° W	Sd	Bond et al. (1997)	Yes	Yes	No?
N. Atlantic	44.5° N	46.33° W	Sd	Bond et al. (1999)	Yes	Yes	No
N. Atlantic	56.37° N	27.81° W	Sd	Bianchi & McCave (1999)	Yes	Yes	No?
N. Ellesmere Island	81° N	80° W	Sd+Lf	Lamoureux & Bradley (1996)	Yes	Yes	–
SW Baltic Sea (Bornholm Basin)	55.38° N	15.4° E	Sd+Is	Andrén et al. (2000)	Yes	Yes	No
N. Fennoscandia (L. Tsuolbmajavri)	68.68° N	22.08° E	Lf	Korhola et al. (2000)	Yes	Yes	No

Table 1 (continued)

Location	Latitude	Longitude	Type	Source	Answer (1)	Answer (2)	Answer (3)
Switzerland (L. Neuchâtel)	47° N	6.55° W	Lf+Is	Filippi et al. (1999)	Yes	Yes	Yes?
NW Scotland (Assynt)	58.11° N	5.06° W	Sp	Proctor et al. (2000)	Yes	Yes	No
W. Ireland	53.53° N	9.93° W	Is	Blackford & Chambers (1995)	Yes	Yes	–
SW Ireland	52.5° N	9.25° W	Sp	McDermott et al. (2001)	Yes	Yes	–
Bermuda Rise	32.17° N	64.5° W	Is	Keigwin (1996)	Yes	Yes	No
Chesapeake Bay	37–38.4° N	76.1° W	Sd	Verardo et al. (1998)	Yes	Yes	–
NW Alaska (Farewell L.)	62.55° N	153.63° W	Lf+Is	Hu et al. (2001)	Yes	Yes	No
S. Dakota (Pickerel L.)	45.51° N	97.27° W	Lf	Dean & Schwalb (2000)	Yes	Yes	No
N. Dakota (Moon L.)	46.85° N	98.16° W	Lf	Laird et al. (1996)	Yes	Yes	No
N. Dakota (Rice L.)	48.01° N	101.53° W	Lf	Yu & Ito (1999)	Yes	Yes	No
Yellowstone P. (Lamar Cave)	44.56° N	110.24° W	Pf+Is	Hadly (1996)	Yes	Yes	No
Colorado Plateau (L. Canyon)	37.42° N	110.67° W	Lf+Gm+Is	Pederson (2000)	Yes	Yes	–
NE Colorado	40–41° N	102–105° W	Gm+Is+D	Madole (1994)	–	Yes	No
SW US (Colorado + Arizona)	34–37.5° N	105–112° W	Lf+Is	Davis (1994)	–	Yes	No?
SW US	32–39° N	109–114° W	Lf+Gm	Ely et al. (1993)	Yes	Yes	No
California (White Mts.)	37.43° N	118.17° W	Is	Feng & Epstein (1994)	Yes	Yes	–
California (L. Owen)	36° N	118.17° W	Is	Li et al. (2000)	Yes	Yes	No
Yucatan Peninsula (L. Chichancanab)	20° N	88.4° W	Lf+Is	Hodell et al. (2001)	Yes	Yes	–
Cariaco Basin	11° N	65° W	Sd	Black et al. (1999)	Yes	Yes	No
Cariaco Basin	10.71° N	65.17° W	Sd+Is	Haug et al. (2001)	Yes	Yes	–
S. Florida	24.95° N	80.55° W	Is	Druffel (1982)	Yes	–	–
SW Puerto Rico	18.12° N	67.09° W	Is	Winter et al. (2000)	Yes	–	–
NE China (Jinchuan)	42.3° N	126.37° E	Is	Hong et al. (2000)	Yes	Yes	No
S. Japan (Yakushima Is.)	30.33° N	130.5° E	Is	Kitagawa & Matsumoto (1995)	Yes	Yes	No
N. India (Pahalgam)	34.02° N	75.20° E	Is	Ramesh (1993)	Yes	–	–
S. India (Nilgiris)	10–10.5° N	77° E	Is	Ramesh (1993)	–	Yes	–
E. Africa (L. Malawi)	10° S	35° E	Lf	Johnson et al. (2001)	Yes	–	No
E. Africa (L. Naivasha)	0.46° S	36.21° E	Lf	Verschuren et al. (2000)	Yes	Yes	No
W. Africa (Cap Blanc)	20.75° N	18.58° W	Is	deMenocal et al. (2000)	Yes	Yes	No
S. Africa	19–35° S	10–33° E	Mp	Tyson & Lindesay (1992)	Yes	Yes	No
S. Africa (Nelson Bay Cave)	34° S	23° E	Is	Cohen & Tyson (1995)	Yes	–	No
S. Africa (Makapansgat)	24.54° S	29.25° E	Sp	Tyson et al. (2000)	Yes	Yes	No
N. New Zealand (Waitomo)	38.27° S	175° E	Sp	Williams et al. (1999)	Yes	–	–
S. New Zealand (Nelson)	40.67° S	172.43° E	Sp	Wilson et al. (1979)	Yes	Yes	No
S. America (several regions)	33–38° S	59.3–67° W	Mp	Iriondo (1999)	Yes	Yes	–
C. Argentina	29.5–35° S	61.75–65.75° W	Gm+D	Carignano (1999)	Yes	Yes	No
C. Argentina	28–36° S	61–67° W	G+Mp	Cioccale (1999)	Yes	Yes	No
NW Argentina	26.5° S	68.09° W	Sd+Is	Valero-Garcés et al. (2000)	Yes	–	–
W. Antarctica (Palmer Deep)	64.86° S	64.21° W	Sd	Domack et al. (2001)	Yes	Yes	No
W. Antarctica (Siple Dome)	81.65° S	148.81° W	Is	Kreutz et al. (1997)	Yes	–	No

[a]Warming or extreme excursion peaked around 1920–1950 before any signifcant anthropogenic CO_2 release to air

[b]Hughes & Diaz concluded that '[our] review indicates that for some areas of the globe (for example, Scandinavia, China, the Sierra Nevada in California, the Canadian Rockies, and Tasmania), temperatures, particularly in summer, appear to have been higher during some parts of this period than those that were to prevail until the most recent decades of the twentieth century. These regional episodes were not strongly synchronous. Evidence from other regions (for example, the Southeast United States, southern Europe along the Mediterranean, and parts of South America) indicates that the climate during that time was little different to that of later times, or that warming, if it occurred, was recorded at a later time than has been assumed. . . . To the extent that glacial retreat is associated with warm summers, the glacial geology evidence would be consistent with a warmer period in A.D. 900–1250 than immediately before or for most of the following seven hundred years.' The main conclusion of Hughes & Diaz (1994) may be in actual agreement with the qualitative classifcation in our paper

[c]Only documentary, historical and archaeological research results, rather than the solar-output model results, of this paper were referred to

[d]For the Dunde ice cap, Thompson et al. (1989) noted that, according to the $\delta^{18}O$ climate proxy, the 1940s, 1950s and 1980s are at least as warm as the Holocene maximum 6000 to 8000 BP. To confrm Thompson et al. (1989) cf. Fig. 6 in Thompson (2000), because the claim that 1930s–1980s are the warmest of the last 6000–8000 yr is not clear from any figure in Thompson et al. (1989). But the main warming of the 1940s–1950s occurred before a significant rise of anthropogenic CO_2 in the air

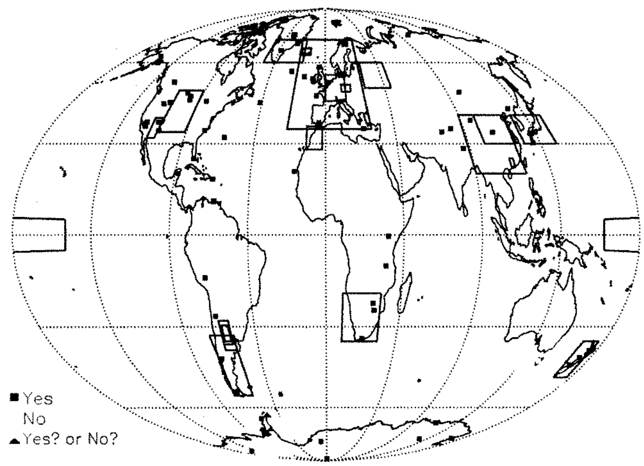


Fig. 1. Geographical distribution of local answers to the following question: Is there an objectively discernible climatic anomaly during the Little Ice Age interval (A.D. 1300–1900) in this proxy record? 'Yes' is indicated by red filled squares or unfilled boxes, 'No' is indicated by green filled circles and 'Yes? or No?' (undecided) is shown with blue filled triangles

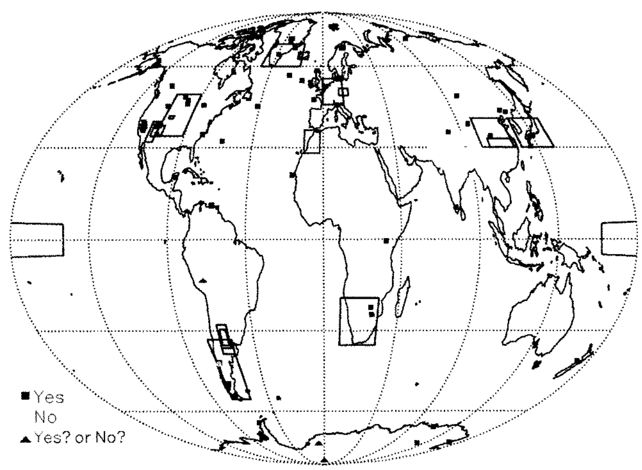


Fig. 2. Geographical distribution of local answers to the following question: Is there an objectively discernible climatic anomaly during the Medieval Warm Period (A.D. 800–1300) in this proxy record? 'Yes' is indicated by red filled squares or unfilled boxes, 'No' is indicated by green filled circles and 'Yes? or No?' (undecided) is shown with blue filled triangles or unfilled boxes

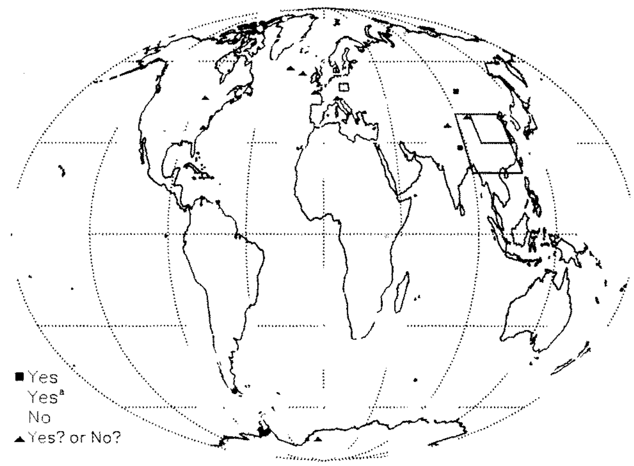


Fig. 3. Geographical distribution of local answers to the following question: Is there an objectively discernible climatic anomaly within the 20th century that is the most extreme (the warmest, if such information is available) period in the record? 'Yes' is indicated by red filled squares, 'No' is indicated by green filled circles or unfilled boxes and 'Yes? or No?' (undecided) is shown with blue filled triangles or unfilled boxes. Answer of 'Yes[a]' is indicated by yellow filled diamonds to mark an early to middle-20th century warming rather than the post-1970s warming

The climate indicators considered here include information from documentary and cultural sources, ice cores, glaciers, boreholes, speleothems, tree-growth limits, lake fossils, mammalian fauna, coral and tree-ring growth, peat cellulose; and pollen, phenological and seafloor sediments. In a rather inhomogeneous way, each proxy is influenced by both climatic and non-climatic factors. We rely on individual researchers for their best judgments in interpreting climatic signals. The 3 questions are addressed in the context of local or regional sensitivity of the proxies to relevant climatic variables, including air temperature, sea surface temperature, precipitation, and any combination of large-scale patterns of pressure, wind and oceanic circulation.

4. UNCERTAINTIES IN INFERRING CLIMATE FROM PROXIES

The accuracy of climate reconstruction from proxies, including the awareness of anthropogenic interventions that could pose serious problems for a qualitative and quantitative paleoclimatology, has been discussed by Bryson (1985), Idso (1989) and others. The temperature changes inferred for the Medieval Warm and Little Ice Age Climatic Anomalies are generally accepted to be no more than 1 to 2°C when averaged over hemispheric or global spatial scale and over decades to a century. Broecker (2001) deduced that only the results from mountain snowline and borehole thermometry are precise to within 0.5°C in revealing changes on a centennial timescale. But the quantification of errors is complex, and both Bradley et al. (2001) and Esper et al. (2002) have challenged Broecker's statement. Earlier, Jones et al. (1998) provided an enlightening review of the quantitative and qualitative limitations of paleoclimatology. Others like Ingram et al. (1978) and Ogilvie & Farmer (1997) had cautioned against the use of quantitative interpretations of climatic results that are based on historical documentation.

In our survey of the literature we have observed 3 distinct types of warnings (Bryson 1985, Clow 1992, Graybill & Idso 1993, Huang et al. 1996, Briffa et al. 1998, Cowling & Skyes 1999, Schleser et al. 1999, Evans et al. 2000, Schmutz et al. 2000, Aykroyd et al. 2001, Ogilvie & Jónsson 2001):

(1) the lack of timescale resolution for the longest-term component of climate signals, e.g. in tree ring and

coral records, or the loss of short-term climate information in borehole temperature reconstructions;

(2) the nonlinearities (related to age, threshold, discontinuous or insufficient sampling, saturated response, limited dynamic range of proxy, etc.) of biological, chemical and physical transfer functions necessary for temperature reconstruction;

(3) the time dependence or nonstationarity of the climate-proxy calibration relations.

Estimates of ground temperature trends from borehole data can be complicated by non-climatic factors associated with changes in pattern of landuse and land cover over time (Lewis & Wang 1998, Skinner & Majorowicz 1999). In general, climate proxies from floral and faunal fossils in lake and bog sediments are only sensitive enough to resolve change to within ±1.3 to 1.8°C (e.g. Lotter et al. 2000). Isotope-coral proxies lack the climate-sensitivity resolution and the continuous length of record to address millennial climatic change. Jones et al. (1998) showed that both coral- and ice-core-based reconstructions performed more poorly than tree-ring records when calibrated against thermometer data since A.D. 1880. On the other hand, tree ring proxies, which usually have annual time resolution, suffer from the loss of information on multidecadal to centennial and longer components of climate change.

The amplitude of large-scale surface temperature change derived from tree-ring proxies can be substantially underestimated—by a factor of 2 to 3 compared to results from borehole thermometry (Huang et al. 2000, Harris & Chapman 2001). It is surprising that the amplitude of climate variability broadly resolved by borehole reconstruction on timescales of at least 50 to 100 yr is larger, rather than smaller, than the high time resolution results from tree-ring proxies, because short-term climate fluctuations are smoothed out by the geothermal heat-flow that acts as a low-pass filter.[1] The different amplitudes found from borehole and tree-ring climate proxies suggest that longer timescale (multi-decades and century) variability is more faithfully captured by borehole results, while the same information can be irretrievably lost from tree-ring records (see e.g. Collins et al. 2002) because of the standardization procedure (to remove bias due to aging of trees). This is why Jones et al. (1998) commented that although one may be confident of intercomparing year-to-year and decade-to-decade (limited to periods shorter than 20 to 30 yr) variability, which should be more sensitively imprinted in tree-ring records, it requires 'considerable faith' to compare, for example, the climate of the 12th and 20th centuries from tree-ring proxies. To date, the practical goal of combining information from borehole and tree-ring proxies, or even comparing borehole and thermometer data, to yield an accurate proxy record that simultaneously resolves timescales of years to centuries, remains unfulfilled.

Despite complicating factors such as the mismatch of climate sensitivities among proxies, a first step has been taken by Beltrami et al. (1995) and Harris & Chapman (2001). Also, Beltrami & Taylor (1995) successfully calibrated a 2000 yr oxygen isotope record from an ice core (near Agassiz) with the help of borehole temperature-depth data (near Neil) for the Canadian Arctic region. Such careful research may help resolve the difficulty of interpreting climate signals that degrade with borehole depth or time. This depth-dependent, increasing degradation has led to the false impression that reconstructed temperatures from geothermal heat flows contained a significantly smaller variability in the distant past than at present.

The approach used here relies on local representations of climate change, which is an advantage because understanding local proxies is the prerequisite for constructing regional and global patterns of change. Another advantage is that by working with a local or regional perspective, we avoid the difficult questions concerning the spatio-temporal coupling of observed changes among various regions and any specific large-scale pattern responsible for those climatic anomalies. Our study has the disadvantage of being non-quantitative. Thus, our assessment falls short of Lamb's (1965) original call for quantitative answers.

An early attempt to study the interlinkage of geographically separated and different proxies, e.g. between marine sediments at Palmer Deep, Antarctica, and atmospheric signals in Greenland ice cores, has been reported by Domack & Mayewski (1999). But many chronologies depend on radiocarbon dating and are too limited in accuracy to allow for reliable interpretation of the timing of events from different areas (e.g. Stine 1998, Domack & Mayewski 1999). The difficult task of areal weighting of different proxy records has been attempted; for the Arctic region by Overpeck et al. (1997), the Northern Hemisphere by Crowley & Lowery (2000), Northern Hemisphere extratropics by Esper et al. (2002) and both Northern Hemisphere and global domains by Mann et al. (1998, 1999, 2000). However, Briffa et al. (2001) criticized the lack of consideration of uncertainties in some of these reconstructions. For example, the composite series in Overpeck et al.'s (1997) reconstruction is not even calibrated with instrumental data.

[1]There are exceptions in careful tree-ring results like those of Esper et al. (2002) that are optimized to capture longer timescale variability

5. RESULTS

Table 1 lists the worldwide proxy climate records we have collected and studied. We restricted the list to records that contain either direct information about the 3 specific questions we posed or at least a continuous time series for 400 to 500 yr. For the majority of cases we followed what individual researchers stated according to their paleoclimatic reconstructions. In a few cases we elaborated on their results in order to remain consistent to our framework.

The figures show the results from Table 1 for the Little Ice Age (Fig. 1), Medieval Warm Period (Fig. 2) and the nature of the 20th century's change (Fig. 3). The figures graphically emphasize the shortage of climatic information extending back to the Medieval Warm Period for at least 7 geographical zones: the Australian and Indian continents, the SE Asian archipelago, large parts of Eastern Europe/Russia, the Middle Eastern deserts, the tropical African and South American lowlands (although the large number of available borehole-heat flow measurements in Australia seems adequate for the reconstruction of ground temperatures back to medieval times; see Huang et al. 2000). Therefore, our conclusions are provisional.

Fig. 1 indicates that Little Ice Age exists as a distinguishable climate anomaly from all regions of the world that have been assessed. Only 2 records—tree ring growth from western Tasmania and isotopic measurements from ice cores at Siple Dome, Antarctica—do not exhibit any persistent climatic change over this period (although the western Tasmania reconstruction contains an exceptionally cold decade centered around 1900; Cook et al. 2000).

Fig. 2 shows the Medieval Warm Period with only 2 negative results. The Himalayan ice core result (Thompson et al. 2000) seems unambiguous, but the tree-ring proxy data from Lenca, southern Chile (Lara & Villalba 1993) is countered by nearby evidence of the Medieval Warm Period (Villalba 1990, 1994).

Fig. 3 shows that most of the proxy records do not suggest the 20th century to be the warmest or the most extreme in their local representations. There are only 3 unambiguous findings favoring the 20th century as the warmest anomaly of the last 1000 yr—the records from the Dyer Plateau, Antarctica, the Himalayas and Mongolia (Thompson et al. 1994, 2000, D'Arrigo et al. 2001). An important, seemingly counter-intuitive, feature of Fig. 3 is the large number of uncertain answers compared to the 2 prior questions, perhaps partly owing to inaccurate calibration between proxy and instrumental data. Also, another feature of the result is the many cases in which the warmest or most extreme climatic anomalies in the proxy indicators occurred in the early to mid-20th century ('Yes[a]'), rather than sustaining throughout the century.

5.1. Glaciers

Broadly, glaciers retreated all over the world during the Medieval Warm Period, with a notable, but minor, re-advance between 1050 and 1150 (Grove & Switsur 1994). Large portions of the world's glaciers, both in the Northern and Southern Hemispheres, advanced during the 1300 to 1900s (Grove 2001b, see also Winkler 2000). The world's small glaciers and tropical glaciers have simultaneously retreated since the 19th century, but some glaciers have advanced (Kaser 1999, Dyurgerov & Meier 2000, D. Evans 2000). Kaser (1999) reemphasized the key role played by atmospheric humidity in controlling the net accumulation and ablation of glaciers by modulating the sublimation and long-wave radiative forcing-feedback budgets in both dry and humid areas. So far, the proposition of the 20th century warming as a natural recovery since the Little Ice Age, together with an amplification by anthropogenic CO_2, is plausible but not definitive (Bradley & Jones 1993, Kreutz et al. 1997, Kaser 1999, Beltrami et al. 2000, Dyurgerov & Meier 2000). On the other hand, D. Evans (2000) discussed the possibility of recent widespread recession of glaciers as a glacio-climatic response to the termination of the Little Ice Age and commented that significant warming phases during interglacials, especially those accompanied by relatively warm winters and cool summers, may lead to the onset of another global glaciation.

Additional proxy records used here reveal that the climatic anomaly patterns known as the Medieval Warm Period (ca. A.D. 800–1300) as well as the Little Ice Age interval (A.D. 1300–1900) occurred across the world. The next 2 subsections describe detailed local changes in the Northern and Southern Hemispheres.

5.2. Northern Hemisphere

A composite reconstruction of summer temperature anomaly assuming a simple, uniform weighting of proxy records by Bradley & Jones (1993) showed that the 1530–1730 interval was the coldest period for the Northern Hemisphere, and the 19th century was the second coldest interval in the last 500 yr.

5.2.1. Western Europe

Cold winters and wet summers prevailed during the Little Ice Age in Switzerland, a location with detailed

and reliable information (Pfister 1995). A careful comparison of the Swiss and central England temperature records (Manley 1974) from 1659–1960 reveals a general correspondence of climatic conditions between the 2 regions. In the Andalusia region of southern Spain, rainfall appears to have alternated between dry and wet century-long spells (wet periods persisted 1590–1649 and 1776–1937; dry periods 1501–1589 and 1650–1775) throughout the Little Ice Age, and with no significant difference to the modern dry period of 1938–1997 (Rodrigo et al. 2000). Enhanced fluvial activity was documented in river basins of North, Western and Central Europe for 1250–1550 and 1750–1900 (A. T. Grove 2001). Over Western Europe, Pfister et al. (1998) concluded that severe winters were less frequent and less extreme during 900–1300 than those during 1300–1900. The mild-winter condition was hypothesized by Pfister et al. (1998) to have caused the northward migration of Mediterranean subtropical plants, where St. Albertus Magnus noted the abundance of pomegranates and fig trees in the 13th century around Cologne and parts of the Rhine valley. Olive trees, which, like fig trees, are also sensitive to prolonged periods of air temperature below freezing, must have grown in Italy (Po valley), France and Germany because a chronicler documented the damage to the olive trees by the bitter frost in January 1234. Lamb (1965) noted generally wet winters but drier summers for the lowlands in England, Ireland, the Netherlands, Denmark, Sweden and NW Germany from about 1200 to 1400. Those conditions are supported by documentary records that describe frequent flooding and storms around those regions during the transitional period between Medieval Warm Period and Little Ice Age.

Is the warmth of the 20th century for western Europe exceptional or unusual? Weather reconstruction results for the Low Countries, the present-day Benelux region, suggest that in order to compare the 20th century to previous centuries, seasonal information in a proxy-climate relation will be required (van Engelen et al. 2001). For example, van Engelen et al. (2001) showed that when the historical reconstructed series from about 800 to 2000 were calibrated to the instrumental temperature records at De Bilt, the 20th century winter temperatures may have been slightly higher (about 0.5°C—the quantitative information on the 20th century warmth is certainly within the margin of uncertainties) than the high winter temperatures of 1000–1100 but that recent warming began in the 19th century. In contrast, the 20th century summer temperatures are neither unusual nor extraordinarily warm when compared to summer temperature variabilities during other times of the 2nd millennium (see Figs. 1 & 2 of van Engelen et al. 2001).

5.2.2. North Atlantic and other oceans

During the Little Ice Age, extensive areas around Mediterranean Europe and the North Atlantic, including Western and Northern Europe, Greenland and Iceland, were experiencing unusually cold and wet conditions as well as many extreme weather events, including deluges, landslides and avalanches (Grove 1996, Ogilvie et al. 2000, A. T. Grove 2001). Climate over Iceland was found, based on various proxies, to be mild from 870 to 1170, with cold periods setting in after 1200. Instead of being a period of unrelenting cold, Ogilvie (1984) emphasized that the most notable aspect of climate over Iceland during the 17th to 19th centuries, with its very cold decades during the 1690s, 1780s, 1810s and 1830s, was its large year-to-year variability (see also Ogilvie & Jónsson 2001).

The viking colonization of Greenland's coastal area starting in 986 is well documented, and the generally mild and benign climatic conditions from about 800 to 1200 that helped to sustain the settlement, are also well supported by ice core and borehole proxy information (Dansgaard et al. 1975, Dahl-Jensen et al. 1998) The Norsemen's 'Western Settlement' (around the Godthab district) was mysteriously abandoned sometime between 1341 and 1362, while the 'Eastern Settlement', actually near the southernmost tip of west Greenland, around the Narssaq and Julianehab districts, died out between 1450 and 1500 (Grove 1996, Ogilvie et al. 2000). It also seems that both cultural and political factors contributed in making the Norse Greenlanders at the Western Settlement more vulnerable to the harsh climatic conditions (Barlow et al. 1997). The timing for the abandonment of these settlements coincided with the general cooling over Greenland, as established by both ice-core isotopic and borehole thermometry (Dansgaard et al. 1975, Stuiver et al. 1995, Dahl-Jensen et al. 1998). From sediment cores near Nansen Fjord, East Greenland, Jennings & Weiner (1996) confirmed an initial cooling between 1270 and 1370, together with the most severe and variable climatic conditions around the East Greenland region from 1630 to 1900. The results of Ogilvie et al. (2000) and Ogilvie & Jónsson (2001) suggest that the overall climatic conditions in the North Atlantic (50–80° N, 0–60° W), especially near Iceland during the 20th century, including the 1970s to 1990s, were neither unusual nor extreme.

In the Mediterranean basin, the island of Crete experienced many severe winters and prolonged droughts during the winter and spring seasons between 1548 and 1648 (Grove & Conterio 1995). In Morocco, the climate during the 16th, 17th and 18th centuries was generally more variable, with frequently drier conditions, than in the early to mid-20th century (Till & Guiot

1990). But no distinctive precipitation anomaly was observed for Morocco during the Medieval Warm Period, although just like conditions during the Little Ice Age, an episode of notable drought occurred from 1186 to 1234. Thus, precipitation anomalies for the Little Ice Age and Medieval Warm Period do not differ from each other in this region—both intervals suffered from persistent drought conditions. For this reason, we gave an uncertain designation for the occurrence of a distinct climatic anomaly associated with Medieval Warm Period for Morocco in Table 1 and Fig. 2.

Distinctly cooler conditions prevailed over the ocean surface—in the Caribbean Sea by about 2 to 3°C (Winter et al. 2000) and in the Sargasso Sea by about 1°C (Keigwin 1996), especially during the 17th and 18th centuries as opposed to the present. DeMenocal et al. (2000) showed that the subtropical North Atlantic's sea surface temperature off Cap Blanc of Mauritania (west Africa) was also cooler by 3 to 4°C between 1300 and 1850 than at present. During the Medieval Warm Period, the Sargasso sea surface temperature was about 1°C warmer than the present-day value (Keigwin 1996), while the sea surface temperature off the coast of Mauritania was only marginally warmer than at present (DeMenocal et al. 2000). High-resolution coral skeletal $\delta^{18}O$ and Sr/Ca ratio records from Bermuda indicate sea-surface temperature standard deviations of about ±0.5°C on interannual and ±0.3°C on decadal timescales during the 16th century, and those ranges of variability are comparable to estimates from modern 20th century's instrumental data (Kuhnert et al. 2002). But those sub-annually resolved coral proxy data also show that although there may be large-scale climate signals like the North Atlantic Oscillation detectable at Bermuda, no correlation can be found with other northern-hemispheric-wide proxy reconstructions because of large spatial differences in climate variability.

From sedimentary concentrations of titanium and iron, Haug et al. (2001) inferred a very dry climate for the Cariaco Basin during the Little Ice Age and relatively wetter conditions in the Medieval Warm Period. Over the equatorial Central Pacific, around the NINO3.4 (5°N–5°S, 160°E–150°W) region, Evans et al. (2000), in their skillful reconstruction of the ENSO-like decadal variability of the NINO3.4 sea surface temperature (SST), showed that there appeared to be a sustained cool phase of the proxy NINO3.4 SST variability from about 1550 to about 1895, hence extending the geographical area covered by the Little Ice Age Climate Anomaly. Evans et al. (2000) also added that the reconstructed NINO3.4 decadal-scale SST variability prior to the 17th century is similar to that of the 20th century, thus suggesting that the recent 20th century decadal-scale changes in the equatorial Pacific Ocean are neither unusual nor unprecedented.

5.2.3. Asia and Eastern Europe

From 49 radiocarbon-dated subfossil wood samples, Hiller et al. (2001) found that the alpine tree-limit on the Khibiny low mountains of the Kola Peninsula was located at least 100 to 140 m above current tree-limit elevation during the relatively warmer time between 1000 and 1300. The summer temperatures corresponding to this tree-line shift during this warm time are estimated to have been at least 0.8°C warmer than today. Based mostly on documentary evidence, Borisenkov (1995) noted that Little Ice Age conditions began as early as the 13th century in Russia with the characteristic of frequent climate extremes both in terms of severe winters, rainy and cool summers, and sustained droughts (up to a decade long). Middle Russia (around 50–60°N and 30–50°E) seems to have experienced the coolest winters around 1620–1680, the coolest summers-springs around 1860–1900, as well as distinctively warm conditions during the first half of the 16th century, similar to conditions for Western Europe described above. The ground surface temperature histories deduced from boreholes around the Czech Republic suggest that winters during 1600 and 1700 were the most severe compared to any other winters since at least 1100 (Bodri & Cermák 1999). The temperature-depth borehole records also yield a clear signature of an anomalously warm period for central Bohemia, especially around 1100–1300.

Bradley and Jones (1993) showed that the mid-17th century was the coldest in China. In NE China, frequent occurrences of extremely dry conditions prevailed during the 16th and 17th centuries (Song 2000). The dry conditions returned again in the 20th century, and now cover a wider area (with indications including the increasing number of days with no discharge from the Yellow River; but these 20th-century events are likely to be confused with other man-made factors). Chan & Shi (2000) further documented the notably larger number of land-falling typhoons over Guandong Province in the early-to-mid-19th century based on a homogeneous set of typhoon records from 1470 to 1931. Using a $\delta^{18}O$ proxy record from peat cellulose with 20 yr resolution and various Chinese historical records, Hong et al. (2000) showed the general cooling trend in the surface air temperature during the Little Ice Age interval in NE China. Hong et al. found 3 of the coolest minima in the record centered around 1550, 1650 and 1750. An obvious warm period peaked around 1100–1200, coinciding with the Medieval Warm Period. The study of documented cultivation of *Citrus reticulata* Blanco (a citrus tree) and *Boehmeria nivea* (a perennial herb), both subtropical and temperature-sensitive plants, during the last 1300 yr showed that northern boundaries for these plants had

shifted and expanded; their northernmost location was reached around 1264 (Zhang 1994). Zhang then deduced that temperature conditions in the 13th century around central China must have been about 1°C warmer than present. Ren (1998) found further evidence from a fossil pollen record at Maili Bog, NE China, that summer monsoon rainfall from 950 to 1270 must have been generally more vigorous in order to explain the high deposition of several pollen taxa, which are otherwise unexplainable by human activity at those times.

Based on the less precise climate proxies like cherry blossom viewing dates, lake freezing dates and historical documentation of climate hazards and unusual weather, Tagami (1993, 1996) found that a warm period prevailed between the 10th and 14th centuries, and a cold period between the late 15th and 19th centuries over large parts of southern Japan. From the study of number of days with snowfall relative to days with rainfall, Tagami (1996) concluded that the 11th and 12th centuries were unusually warm in Japan. During the Little Ice Age, summers were relatively cool from the 1730s to 1750s, in the 1780s, from the 1830s to the 1840s and in the 1860s, and winters were cold through the 1680s to 1690s, and in the 1730s and 1810s. From the tree-cellulose $\delta^{13}C$ record of a giant Japanese cedar *Cryptomeria japonica* grown on Yakushima Island of southern Japan, Kitagawa & Matsumoto (1995) inferred a cool temperature of 2°C below average from 1600 to 1700 and a warm period of about 1°C above average between 800 and 1200.

5.2.4. North America

Overall, the composite summer temperature anomaly from Bradley & Jones (1993) showed that over North America the temperature during the 15th to 17th centuries was 1°C cooler than the average of the reference period 1860–1959. Over the southern Sierra Nevada, California, Graumlich (1993) found that the coolest 50 yr interval in the 1000 yr tree-ring proxy record was around 1595–1644, while the wettest 50 yr period was 1712–1761. These periods are consistent with our definition of a discernible climatic anomaly associated with the Little Ice Age interval of 1300–1900. Ely et al. (1993) noted from river records in Arizona and Utah that the most extreme flooding events occurred during transitions from cool to warm climate conditions, especially during the late 1800s to early 1900s. For the Central U.S.A. (33–49° N, 91–109° W), drought episodes were noted for the 13th to 16th centuries (Woodhouse & Overpeck 1998). These droughts were of longer duration and greater extent than the 1930s–1950s drought. Additionally, both Yu & Ito (1999) and Dean & Schwalb (2000) pointed to the cycles of aridity lasting about 400 yr from lake records of the Northern Great Plains, where the last dry condition peaked around 1550–1700.

From an extensive collection of multiproxy evidence, Stine (1998) concluded that during the Medieval Warm Period prolonged intervals of extreme drought visited California, the NW Great Basin, the northern Rocky Mountains/Great Plains, while markedly wetter regimes persisted in the upper Midwest/sub-Arctic Canada and Southern Alaska/British Columbia areas. There was also a significant but brief interval around 1110–1140 when moisture conditions switched from dry to wet in California, the NW Great Basin, the northern Rocky Mountains/Great Plains, and from wet to dry in the upper Midwest/sub-Arctic Canada and southern Alaska/British Columbia. The most likely explanation for this rapid and dramatic switch from wet to dry conditions in the upper Midwestern U.S. around 1100 is the contraction and subsequent expansion of the circumpolar vortex. Summer polar fronts shifted significantly southward, and stopped the penetration of moisture-laden air from the Gulf of Mexico (Bryson et al. 1965). Stine (1998) added the requirement of a concomitant jet-stream change, from zonal to azonal, in order to explain the distinct observed differences of the moisture patterns between the upper Midwest and southern Alaska/British Columbia. Graumlich (1993)'s reconstruction of summer temperature and winter precipitation from trees in the Sierra Nevada confirmed the overall warm and dry conditions for California during medieval times, when the 2 warmest and the 2 driest 50 yr intervals occurred, at 1118–1167, 1245–1294 and 1250–1299, 1315–1364, respectively.

Hu et al. (2001), based on their high-resolution (multidecadal) geochemical analysis of sediments from Farewell Lake by the NW foothills of the Alaska Range, also found pronounced signatures of the Medieval Warm Period around 850–1200. During the Little Ice Age the surface water temperature of Farewell Lake suffered a low in 1700 calibrated to be about 1.75°C cooler than present. They also noted that colder periods were in general wetter (in contrast to drier conditions during Little Ice Age in the Central U.S. region described above) than the warm periods in this part of NW Alaska. On the Yucatan Peninsula, prolonged drought episodes recur approximately every 208 yr, with the 2 most significant recent peaks centered around 800 and 1020 (Hodell et al. 2001). The timing of severe droughts also seems to fit several known discontinuities in the evolution of the Mayan culture.

5.3. Southern Hemisphere

Figs 1–3 highlight the sparse coverage of the Southern Hemisphere by proxy climatic information through the 2nd millennium.

5.3.1. New Zealand

In New Zealand, the $\delta^{18}O$ concentration in a stalagmite record from a cave in NW Nelson showed the coldest times during the Little Ice Age to be around 1600–1700, while exceptionally warm temperatures occurred around 1200–1400, in association with the general phenomenology of the Medieval Warm Period (Wilson et al. 1979). The cooling anomaly around 1600–1700 apparent in the $\delta^{18}O$ stalagmite record coincides with the smallest growth ring (i.e. coolest period) for the silver pine *Lagarostrobus colensoi* from Mangawhero of the North Island. However, at Ahaura in the South Island, the smallest ring width index of the 600 yr record occurred about 1500–1550 (D'Arrigo et al. 1998). Williams et al. (1999) advise caution in interpreting stable isotope data from New Zealand, especially the correctional functional relations among temperature, precipitation and $\delta^{18}O$ data (which are strongly influenced by oceans surrounding New Zealand) from Waitomo, North Island's speleothems. The mean annual temperatures at Waitomo from 1430–1670 were deduced, based on the analysis of $\delta^{18}O$ data from Max's cave stalagmite, to be about 0.8°C cooler than today.

5.3.2. South Africa

Tyson et al. (2000) showed through isotopic measurements in speleothem that the interior region of South Africa, near the Makapansgat Valley (eastern part of South Africa), had a maximum cooling of about 1°C around 1700 compared to the present. This cooling feature corresponds well with the maximum cooling signal contained in a coral record from SW Madagascar (Tyson et al. 2000). Tyson & Lindesay (1992) showed that the Little Ice Age in South Africa had 2 major cooling phases, around 1300–1500 and 1675–1850, with a sudden warming interval between 1500 and 1675. In addition, Tyson & Lindesay suggested a weakening of the tropical easterlies that increased the incidence of drought during the Little Ice Age in South Africa—with a relatively drier condition for the summer rainfall region in the northeast, but a wetter condition for the winter rainfall region near the coastal Mediterranean zone in the southwest. At Makapansgat Valley, the Medieval Warm Period peaked with a temperature about 3–4°C warmer than present around 1200–1300 (Tyson et al. 2000). The multiproxy review by Tyson & Lindesay (1992) showed evidence for a wetter South Africa after 1000, when forest and wetland become more extensive, including the development of a riverine forest in the northern Namib desert along the Hoanib river during the 11th–13th centuries.

5.3.3. South America

Over southern South America's Patagonia, the Little Ice Age's climatic anomalies, as deduced from tree ring records, were manifest as cold and moist summers with the most notable, persistent century-long wet intervals centered around 1340 and 1610 (Villalba 1994). From a multiproxy study of lacustrine sediments at Lake Aculeo (about 34° S; 50 km southeast of Santiago, Chile), Jenny et al. (2002) found a period of greatly increased flood events centered around 1400–1600 (and in 3 other intervals: 200–400, 500–700 and 1850–1998), which could be interpreted as increased winter rains from enhanced mid-latitude westerlies that ushered in more frontal system activities. In contrast, during the Medieval Warm Period, the southern Patagonia region at latitudes between 47 and 51°S became abnormally dry for several centuries before 1130 when water levels in several lakes (Lake Argentino, Lake Cardiel and Lake Ghio) dropped significantly. Also, trees like the southern beech *Nothofagus* sp. grew as old as 100 yr in the basin of these lakes before being killed by reflooding of the lakes (Stine 1994).

Slightly north toward the central region of Argentina (around Córdoba Province), Carignano (1999), Cioccale (1999) and Iriondo (1999) noted the prevailing conditions for the advancement of the Andean glaciers during the Little Ice Age, with 2 distinct cold and dry intervals around the 15th to 16th, and the 18th to the early 19th centuries. The significant climate aridification and deterioration in central Argentina (in contrast to the more humid conditions and increased flood frequency in central Chile near Lake Aculeo) during the Little Ice Age interval is supported by the formation of large, parabolic sand dunes 150–200 m long, 60–80 m wide, and 2–3 m high in the Salinas Grandes basin (Carignano 1999). Meanwhile the Mar Chiquita Lake was transformed into a swamp surrounded by dunes in the 18th century. Today Mar Chiquita is the largest lake in Argentina, covering a surface area of 6000 km^2 and with a depth of 13 m (Iriondo 1999). The climatic conditions during the Medieval Warm Period around Central Argentina were generally warmer and more humid than at other times in the 2nd millennium, when the dune fields were conquered by lakes and the Mar Chiquita Lake expanded beyond its present dimensions. Precipitation exceeded current val-

ues, and the mean local temperature may have been about 2.5°C warmer, perhaps because of the southward shift of the tropical climate belt into this area (Iriondo 1999). The northern part of Córdoba Province was invaded by the eastern boundary of the Chaco Forest, which is located hundreds of kms to the northwest today (Carignano 1999). Cioccale (1999) noted evidence of human cultivation of hillside areas in Central Andes, Peru, at places as high as 4300 m above sea level around 1000.

5.3.4. Antarctica

The last important source of geographical information for conditions during the Medieval Warm Period and the Little Ice Age in the southern hemisphere is obtained from glaciers, ice cores and sea sediments studies on and around Antarctica. Although many notable physical, biological and environmental changes have recently occurred there, especially around the Antarctic Peninsula during the last 50 yr (Mercer 1978, Thomas et al. 1979, Rott et al. 1996, Vaughan & Doake 1996, Smith et al. 1999, Doran et al. 2002, Marshall et al. 2002), most of the 20th-century changes contained in the proxy records discussed here cannot be considered extreme or unusual (see Fig. 3, also Vaughan & Doake 1996, D. Evans 2000).

For the Little Ice Age, advances of glaciers in South Georgia Island, which is currently half-covered by glaciers, began after the late 13th century, with a peak advancement around the 18th–20th centuries (Clapperton et al. 1989). Glacier retreats occurred after about 1000, which corresponds to the timing for the Medieval Warm Period. Baroni & Orombelli (1994) described a similar sequence for glacier advances and retreats during the Little Ice Age and Medieval Warm Period for the Edmonson Point glacier at the Terra Nova Bay area of Victoria Land on the Antarctic continent (East Antarctica). The Edmonson Point glacier retreated in 2 distinct phases, around 920–1020 and 1270–1400, and then advanced at least 150 m after the 15th century. Isotopic thermometry from ice cores at Dome C (74.65° S, 124.17° E, elevation 3240 m) and Law Dome (66.73° S, 112.83° E, elevation 1390 m) both indicate cooler and warmer anomalies for the Little Ice Age and Medieval Warm Period respectively (Benoist et al. 1982, Morgan 1985). High-resolution records of magnetic susceptibility from deep sea cores (Domack & Mayewski 1999, Domack et al. 2001) drilled near Palmer Deep site (64.86° S, 64.21° W) off the Antarctic Peninsula also show a marked increased, in bioproductivity, hence a decrease in magnetic susceptibility because of dilution of the magnetite, with a peak centered around 1000–1100 yr BP. This observation probably implied warm temperatures and minimal sea-ice conditions, coinciding with the Medieval Warm Period. In the same record, Domack and colleagues found a decrease in bio-productivity, hence an increase in magnetic susceptibility owing to less dilution of the magnetic minerals by biogenic materials, from about 700 to 100 yr BP. This time period corresponds to the Little Ice Age of ca. 14th–19th centuries and is likely to have been accompanied by cool and windy conditions. Abundance analyses of Na+ sea salt in the ice core from Siple Dome (81.65° S, 148.81° W) also confirm the Little Ice Age anomaly characterized by substantial variability in the strength of meridional circulation that extended between 1400 and 1900 (Kreutz et al. 1997).

But there are also indications for significant regional differences in climatic anomalies associated with the 2 phenomena at Antarctica. The temperature at Siple Station (75.92° S, 84.25° W, elevation 1054 m) was relatively warm from the 15th to early 19th centuries (although there were also noticeable decade-long cooling dips centered around 1525, 1600 and 1750; Mosley-Thompson 1995). The 400 yr isotopic temperature inferred from a core at the Dalinger Dome (64.22° S, 57.68° W, elevation 1640 m) on James Ross - Island, off the Antarctic Peninsula, also showed 1750–1850 to be the warmest interval, followed by a cooling of about 2°C since 1850 and continuing to 1980 (Aristarain et al. 1990). A recent borehole temperature reconstruction from Taylor Dome, east Antarctica (77.8° S, 158.72° E, elevation 2374 m), also reported the same inverted temperature anomalies, during which the Little Ice Age interval was about 2°C warmer, while the coldest temperature of the past 4000 yr was reached around 1000 (Clow & Waddington 1999); note that we have omitted these discussions in our Table 1 or Figs. 1 to 3 because the results are claimed as preliminary and they were only presented in a conference abstract.

6. DISCUSSION

The widespread geographical evidence assembled here supports the existence of both the Little Ice Age and the Medieval Warm Period, and should serve as useful validation targets for any reconstruction of global climate history of the last 1000 yr. Our results suggest a different interpretation of the multiproxy climates compared to recent conclusions of Mann et al. (1998, 1999, 2000). Because the calibration of proxy indicators to instrumental data is still a matter of open-ended research (with differing sensitivities not only for the same proxy at different locations but also for different proxies at the same location), it is premature to select a year or decade as the warmest or coldest in a multiproxy-based record.

Barnett et al. (1999) has pointed out that it is *impossible* to use available instrumental records to provide estimates for the multi-decadal and century-long type of *natural* climatic variations, owing to the specific period and short duration of instrumental records. Thus, paleo-proxies remain the only hope for assessing the amplitude and pattern of climatic and environmental change in the pre-anthropogenic era. We agree with Barnett et al. (1999) that each proxy should be studied first in terms of local change before several records can be combined for regional and larger spatial-scale analyses and interpretations. The conclusion derives mainly from the real possibility of non-stationarity in the proxy-climate calibration to instrumental records, the lack of adequate superposition rules given variability in each type of proxy, as well as the lack of clear physical understanding on the multidecadal climate variability from theoretical or empirical studies.

All current calibration of proxies to large-scale instrumental measurements have been mainly valid over phases of rising temperature (Ogilvie & Jónsson 2001). The concern is that a different calibration response arises when the procedure is extended to an untested climate regime associated with a persistent cooling phase. Evans et al. (2002) worried about the reality of spurious frequency evolution that may contaminate a multiproxy reconstruction, in which the type of proxy data changes over time and no sufficient overlap of proxy data exists for a proper inter-proxy calibration/validation procedure. In other words, each proxy may have its distinct frequency response function, which could confuse the interpretation of climate variability. Finally, another concern is the lack of understanding of the air-sea relationship at the multidecadal timescale, even in the reasonably well observed region of the North Atlantic (Häkkinen 2000, Seager et al. 2000, Marshall et al. 2001, Slonosky & Yiou 2001, von Storch et al. 2001).

Briffa (2000) concluded that dendroclimatological records may support 'the notion that the last 100 years have been unusually warm, at least within the context of the last two millennia' Slightly later, Briffa et al. (2001), by adopting a new analysis procedure that seeks to preserve greater, long timescale variability (which shows a notable increase in variance at the 24–37 yr timescale compared with a previous standardization procedure) in their tree-ring density data than previously possible, stated that the 20th century is the globally warmest century of the last 600 yr. This conclusion is consistent with the borehole reconstruction results of Huang et al. (2000). However, longer and more carefully reconstructed tree-ring chronologies from Esper et al. (2002) showed that the Medieval Warm Period was as warm as the 20th century for at least a region covering the Northern Hemisphere extratropics from about 30 to 70° N.

An important aspect of both the Briffa et al. (2001) and Esper et al. (2002) studies is the new derivation of formal, time-dependent standard errors for their temperature reconstructions, amounting to about ±0.1 to 0.3°C from 1000 through 1960 (see also Jones et al. 1999, 2001). This assignment of standard errors contrasts with those assigned in Mann et al.'s (1999) annually-resolved series, where the uncertainties were assigned only for pre-instrumental data points in their original publication (that assumption of 'error-free' instrumental thermometer data is incorrect—see Jones et al. 1999, Folland et al. 2001). Over the full 2nd millennium, Esper et al. (2002) deduced a slightly larger range in their confidence limits after 1950 (compared to the pre-1950 interval extending back to 800) and attributed those higher uncertainties to the accounting for the anomalous modern ring-growth problem (Graybill & Idso 1993, Jacoby & D'Arrigo 1995, Briffa et al. 1998, Feng 1999, Barber et al. 2000, Jacoby et al. 2000, Knapp et al. 2001).

7. CONCLUSION

Many interesting questions on the geographical nature and physical factors of surface temperature or precipitation changes over the last 1000 yr cannot be quantitatively and conclusively answered by current knowledge. The adopted period of 1000 yr is strictly a convenience that merits little scientific weight.

Climate proxy research provides an aggregate, broad perspective on questions regarding the reality of Little Ice Age, Medieval Warm Period and the 20th century surface thermometer global warming. The picture emerges from many localities that both the Little Ice Age and Medieval Warm epoch are widespread and near-synchronous phenomena, as conceived by Bryson et al. (1963), Lamb (1965) and numerous researchers since. Overall, the 20th century does not contain the warmest anomaly of the past millennium in most of the proxy records, which have been sampled world-wide. Past researchers implied that unusual 20th century warming means a global human impact. However, the proxies show that the 20th century is not unusually warm or extreme.

The lack of unusual warmth in the 20th century does not argue against human impacts on local and regional scales (perhaps on scales as small as 10 to 1000 km^2 for precipitation and 10^4 to 10^5 km^2 for temperature). Recently, Lawton et al. (2001) demonstrated how the deforested areas of tropical lowlands can, in combination with favorable topographical conditions and altered atmospheric air flow across the landscape, significantly raise the bases of convective and orographic

clouds around the Monteverde montane cloud forests of Costa Rica during the dry season, and thus drastically impact local ecosystems. It is thus clear that human activity affects local environment and climate. On the other hand, A. T. Grove (2001) advised caution when interpreting dubious claims about the dominant role played by human activity such as in deforestation, agricultural expansion and population growth on geomorphological changes in Mediterranean Europe. In particular, A. T. Grove (2001) showed that the wide spread occurrence of the medieval fluvial terrace called the 'younger fill' around Mediterranean Europe was more powerfully influenced by increased frequency of deluges during Alpine glacier advances associated with the Little Ice Age than by soil erosion and rapid sedimentation caused by deforestation.

Acknowledgements. This work was supported by funds from the American Petroleum Institute (01-0000-4579), the Air Force Office of Scientific Research (Grant AF49620-02-1-0194) and the National Aeronautics and Space Administration (Grant NAG5-7635). The views expressed herein are those of the authors and are independent of the sponsoring agencies. We have benefitted greatly from the true and kind spirit of research communications (including a preview of their thoughts) with the late Jean Grove (who passed away on January 18, 2001), Dave Evans, Shaopeng Huang, Jim Kennett, Yoshio Tagami and Referee #3. We thank John Daly, Diane Douglas-Dalziel, Craig and Keith Idso for their unselfish contributions to the references. We also thank the Editor, Chris de Freitas, for very helpful editorial changes that improved the manuscript. We are very grateful to Maria McEachern, Melissa Hilbert, Barbara Palmer and Will Graves for invaluable library help, and both Philip Gonzalez and Lisa Linarte for crucial all-around help.

LITERATURE CITED

Andrén E, Andrén T, Sohlenius G (2000) The Holocene history of the southwestern Baltic Sea as reflected in sediment core from the Bornholm Basin. Boreas 29:233–250

Aristarain AJ, Jouzel J, Lorius C (1990) A 400 years isotope record of the Antarctic Peninsula climate. Geophys Res Lett 17:2369–2372

Arseneault D, Payette S (1997) Reconstruction of millennial forest dynamics from tree remains in a subarctic tree line peatland. Ecology 78:1873–1883

Aykroyd RG, Lucy D, Pollard AM, Carter AHC, Robertson I (2001) Temporal variability in the strength of the proxy-climate correlations. Geophys Res Lett 28:1559–1562

Barber VA, Juday GP, Finney BP (2000) Reduced growth of Alaskan white spruce in the twentieth century from temperature-induced drought stress. Nature 405:668–673

Barclay DJ, Wiles GC, Calkin PE (1999) A 1119-year tree-ring-width chronology from western Prince William Sound, southern Alaska. Holocene 9:79–84

Barlow LK, Sadler JP, Ogilvie AEJ, Buckland PC and 5 others (1997) Interdisciplinary investigations of the end of the Norse Western Settlement in Greenland. Holocene 7:489–499

Barnett T, Hasselmann K, Chelliah M, Delworth T and 7 others (1999) Detection and attribution of recent climate change: a status report. Bull Am Meteor Soc 80:2631–2659

Baroni C, Orombelli G (1994) Holocene glacier variations in the Terra Nova Bay area (Victoria Land, Antarctica). Antarctic Sci 6:497–505

Beltrami H, Taylor AE (1995) Records of climatic change in the Canadian Arctic: towards calibrating oxygen isotope data with geothermal data. Glob Planet Change 11:127–138

Beltrami H, Chapman DS, Archambault S, Bergeron Y (1995) Reconstruction of high resolution ground temperature histories combining dendrochronological and geothermal data. Earth Planet Sci Lett 136:437–445

Beltrami H, Wang J, Bras RL (2000) Energy balance at the Earth's surface: heat flux history in eastern Canada. Geophys Res Lett 27:3385–3388

Benoist JP, Jouzel J, Lorius C, Merlivat L, Pourchet M (1982) Isotope climatic record over the last 2.5 ka from Dome C, Antarctica, ice cores. Ann Glaciology 3:17–22

Bernabo JC (1981) Quantitative estimates of temperature changes over the last 2700 years in Michigan based on pollen data. Quat Res 15:143–159

Bianchi GG, McCave IN (1999) Holocene periodicity in North Atlantic climate and deep-ocean flow south of Iceland. Nature 397 515–517

Biondi F, Perkins DL, Cayan DR, Hughes MK (1999) July temperature during the second millennium reconstructed from Idaho tree rings. Geophys Res Lett 26: 1445–1448

Black DE, Peterson LC, Overpeck JT, Kaplan A, Evans MN, Kashgarian M (1999) Eight centuries of North Atlantic Ocean atmosphere variability. Science 286: 1709–1713

Blackford JJ, Chambers FM (1995) Proxy climate record for the last 1000 years from Irish blanket peat and a possible link to solar variability. Earth Planet Sci Lett 133:145–150

Bodri L, Cermák (1999) Climate change of the last millennium inferred from borehole temperatures: regional patterns of climatic changes in the Czech Republic. Part III. Glob Planet Change 21:225–235

Bond G, Showers W, Cheseby M, Lotti R and 6 others (1997) A pervasive millennial-scale cycle in North Atlantic holocene and glacial climates. Science 278:1257–1266

Bond G, Showers W, Elliot M, Evans M, Lotti R, Hajdas I, Bonani G, Johnson S (1999) The North Atlantic's 1–2 kyr climate rhythm: relation to Heinrich events, Dansgaard/Oeschger cycles and the little ice age. In Clark PU, Webb RS, Keigwin LD (eds) Mechanisms of global climate change at millennial time scales, Vol 112. American Geophysical Union Press, Washington DC, p 35–58

Borisenkov YeP (1995) Documentary evidence from the U.S.S.R. In: Bradley RS, Jones PD (eds) Climate since A.D. 1500. Routledge, New York, p 171–183

Bradley RS, Jones PD (1993) 'Little Ice Age' summer temperature variations: their nature and relevance to recent global warming trends. Holocene 3:367–376

Bradley RS, Briffa KR, Crowley TJ, Hughes MK, Jones PD, Mann ME (2001) The scope of Medieval Warming. Science 292:2011–2012

Briffa KR (2000) Annual climate variability in the Holocene: interpreting the message of ancient trees. Quat Sci Rev 19: 87–105

Briffa KR, Jones PD, Bartholin TS, Eckstein D, Schweingruber FH, Karlen W, Zetterberg P, Eronen M (1992) Fennoscandian summers from AD 500: temperature changes on short and long timescales. Clim Dyn 7:111–119

Briffa KR, Schweingruber FH, Jones PD, Osborn TJ, Harris IC, Shiyatov SG, Vaganov EA, Grudd H (1998) Trees tell of past climates: but are they speaking less clearly today? Phil Trans R Soc Lond B 353:65–73

Briffa KR, Osborn TJ, Schweingruber FH, Harris IC, Jones PD, Shiyatov SG, Vaganov EA (2001) Low-frequency temperature variations from a northern tree-ring density network. J Geophys Res 106:2929–2941

Broecker WS (2001) Was the medieval warm period global? Science 291:1497–1499

Bryson RA (1985) On climatic analogs in paleoclimatic reconstruction. Quat Res 23:275–286

Bryson RA, Bryson RU (1997) High resolution simulations of regional holocene climate: North Africa and the Near East. In: Dalfes HN, Kukla G, Weiss H (eds) Third millenium BC climate change and the Old World collapse. NATO ASI Series, Vol. 149. Springer-Verlag, Heidelberg, p 565–593

Bryson RA, Arakawa H, Aschmann HH, Baerris DA and 36 others (1963) NCAR Technical Note In: Bryson RA, Julian PR (eds) Proc Conf Climate of the 11th and 16th Centuries, Aspen CO, June 16–24 1962, National Center for Atmospheric Research Technical Notes 63-1, Boulder CO

Bryson RA, Irving WN, Larsen JA (1965) Radiocarbon and soil evidence of former forest in the Southern Canadian Tundra. Science 147:46–48

Calkin PE, Wiles GC, Barclay DJ (2001) Holocene coastal glaciation of Alaska. Quat Sci Rev 20:449–461

Carignano CA (1999) Late Pleistocene to recent climatic chance in Córdoba Province, Argentina: geomorphological evidence. Quat Int 57-58:117–134

Chan JCL, Shi JE (2000) Frequency of typhoon landfall over Guangdong province of China during the period 1470–1931. Int J Clim 20:183–190

Cioccale MA (1999) Climatic fluctuations in the central region of Argentina in the last 1000 years. Quat Int 62:35–47

Clapperton CM, Sugden DE, Birnie J, Wilson MJ (1989) Lateglacial and holocene glacier fluctuations and environmental change on South Georgia, Southern Ocean. Quat Res 31:210–228

Clow GD (1992) The extent of temporal smearing in surface-temperature histories derived from borehole temperature measurements. Palaeogeogr Palaeoclim Palaeoecol 98 81–86

Clow GD, Waddington ED (1999) Quantification of natural climate variability in Central Greenland and East Antarctica using borehole paleothermometry (abstract). International Union of Geodesy and Geophysics XXII General Assembly Abstract Book: B249, Birmingham UK

Cohen AL, Tyson PD (1995) Sea-surface temperature fluctuations during the Holocene off the south coast of Africa: implications for terrestrial climate and rainfall. Holocene 5: 304–312

Collins M, Osborn TJ, Tett SFB, Briffa KR, Schweingruber FH (2002) A comparison of the variability of a climate model with paleotemperature estimates from a network of tree-ring indices. J Clim 15:1497–1515

Cook ER, Buckley BM, D'Arrigo RD, Peterson MJ (2000) Warm-season temperatures since 1600 BC reconstructed from Tasmanian tree rings and their relationship to large-scale sea surface temperature anomalies. Clim Dyn 16: 79–91

Cowling SA, Skyes MT (1999) Physiological significance of low atmospheric CO_2 for plant-climate interactions. Quat Res 52:237–242 (See additional research exchanges in Quat Res 53:402–404 and Quat Res 53:405–406)

Crowley TJ, Lowery TS (2000) How warm was the medieval warm period? Ambio 29:51–54

Dahl-Jensen D, Mosegaard K, Gundestrup N, Glow GD, Johnsen SJ, Hansen AW, Balling N (1998) Past temperatures directly from the Greenland ice sheet. Science 282:268–271

Dansgaard W, Johnsen SJ, Reeh N, Gundestrup N, Clausen HB, Hammer CU (1975) Climatic changes, Norsemen and modern man. Nature 255:24–28

D'Arrigo RD, Cook ER, Salinger J, Palmer J, Buckley BM, Krusic PJ, Villalba R (1998) Tree-ring records from New Zealand: long-term context for recent warming trend. Clim Dyn 14:191–199

D'Arrigo RD, Jacoby G, Frank D, Pederson N and 5 others (2001) 1738 years of Mongolian temperature variability inferred from a tree-ring width chronology of Siberian Pine. Geophys Res Lett 28:543–546

Davis OK (1994) The correlation of summer precipitation in the southwestern USA with isotopic records of solar activity during the Medieval Warm Period. Clim Change 26: 271–287

Dean WE, Schwalb A (2000) Holocene environmental and climatic change in the Northern Great Plains as recorded in the geochemistry of sediments in Pickerel Lake, South Dakota. Quat Int 67:5–20

deMenocal P (2001) Cultural responses to climate change during the late Holocene. Science 292:667–673

deMenocal P, Ortiz J, Guilderson T, Sarnthein M (2000) Coherent high- and low-latitude climate variability during the Holocene warm period. Science 288:2198–2202

Domack EW, Mayewski PA (1999) Bi-polar ocean linkages: evidence from late-holocene Antarctic marine and Greenland ice-core records. Holocene 9:247–251

Domack E, Leventer A, Dunbar R, Taylor F, Brackfeld S, Sjunneskog C plus ODP Leg 178 Team (2001) Chronology of the Palmer Deep site, Antarctic Peninsula: a holocene palaeoenvironmental reference for the circum-Antarctic. Holocene 11:1–9

Doran PT, Priscu JC, Lyons WB, Walsh JE and 9 others (2002) Antarctic climate cooling and terrestrial ecosystem response. Nature 415:517–520 (see also exchanges between Turner et al. and Walsh, Doran et al. (2002) in Recent temperature trends in the Antarctic. Nature 418:291–292)

Druffel EM (1982) Banded corals: Changes in Oceanic carbon-14 during the Little Ice Age. Science 218:13–19

Dyurgerov MB, Meier MF (2000) Twentieth century climate change: evidence from small glaciers. Proc Natl Acad Sci USA 97:1406–1411

Ely LL, Enzel Y, Baker VR, Cayan DR (1993) A 5000-year record of extreme floods and climate change in the southwestern United States. Science 262:410–412

Esper J, Cook ER, Schweingruber FH (2002) Low-frequency signals in long tree-ring chronologies for reconstructing past temperature variability. Science 295:2250–2253

Evans DJA (2000) Ice in a greenhouse: the impacts of global warming on world glacier cover. Trans Inst Engineers and Shipbuilders in Scotland 143:11–19

Evans MN, Kaplan A, Cane MA, Villalba R (2000) Globality and optimality in climate field reconstructions from proxy data. In: Markgraf V (ed) Inter-hemispheric climate linkages. Academic Press, San Diego, CA p 53–72

Evans MN, Kaplan A, Cane MA (2002) Pacific sea surface temperature field reconstruction from coral $\delta^{18}O$ data using reduced space objective analysis. Paleocean 17:10.1029/2000PA000590

Feng X (1999) Trends in intrinsic water-use efficiency of natural trees for the past 100–200 years. Geochimica et Cosmochim Acta 63:1891–1903

Feng X, Epstein S (1994) Climatic implications of an 8000-year hydrogen isotope time series from Bristlecone Pine trees. Science 265:1079–1081

Filippi ML, Lambert P, Hunziker J, Kübler B, Bernasconi S

(1999) Climatic and anthropogenic influence on the stable isotope record from bulk carbonates and ostracodes in Lake Neuchâtel, Switzerland, during the last two millenia. J Paleolim 21:19–34
Folland CK, Rayner NA, Brown SJ, Smith TM and 7 others (2001) Global temperature change and its uncertainties since 1861. Geophys Res Lett 28:2621–2624
Graumlich LJ (1993) A 1000-year record of temperature and precipitation in the Sierra Nevada. Quat Res 39:249–255
Graybill DA, Idso SB (1993) Detecting the aerial fertilization effect of atmospheric CO_2 enrichment in tree-ring chronologies. Glob BioGeoChem Cyc 7:81–95
Grissino-Mayer HD (1996) A 2129-year reconstruction of precipitation for northwestern New Mexico, USA. In Dean JS, Meko DM, Swetnam TW (eds) Tree rings, environment and humanity , Radiocarbon (U Arizona Press), Tucson, p 191–204
Grove AT (2001) The 'Little Ice Age' and its geomorphological consequences in Mediterranean Europe. Clim Change 48: 121–136
Grove JM (1996) The century time-scale. In: Driver TS, Chapman GP (eds) Time-scales and environmental change. Routledge, London p 39–87
Grove JM (2001a) The initiation of the Little Ice Age in regions round the North Atlantic. Clim Change 48:53–82
Grove JM (2001b) The onset of the Little Ice Age. In: Jones PD, Ogilvie AEJ, Davies TD, Briffa KR (eds) History and climate—memories of the future? Kluwer Academic Publishers, New York p 153–185
Grove JM, Conterio A (1995) The climate of Crete in the sixteenth and seventeenth centuries. Clim Change 30: 223–247
Grove JM, Switsur R (1994) Glacial geological evidence for the medieval warm period. Clim Change 26:143–169
Hadly EA (1996) Influence of late-Holocene climate on Northern Rocky Mountain mammals. Quat Res 46:289–310
Häkkinen S (2000) Decadal air-sea interaction in the North Atlantic based on observations and modeling results. J Clim 13:1195–1219
Hansen J, Ruedy R, Glascoe J, Sato M (1999) GISS analysis of surface temperature change. J Geophys Res 104: 30997–31022
Harris RN, Chapman DS (2001) Mid-latitude (30° to 60° N) climatic warming inferred by combining borehole temperatures with surface air temperatures. Geophys Res Lett 28:747–750
Haug GH, Hughen KA, Sigman DM, Peterson LC, Röhk U (2001) Southward migration of the intertropical convergence zone through the holocene. Science 293:1304–1308
Hendy EJ, Gagan MK, Alibert CA, McCulloch MT, Lough JM, Isdale PJ (2002) Abrupt decrease in tropical Pacific sea surface salinity at end of Little Ice Age. Science 295:1511–1514
Hiller A, Boettger T, Kremenetski C (2001) Mediaeval climatic warming recorded by radiocarbon date alpine tree-line shift on the Kola Peninsula, Russia. Holocene 11: 491–497
Hodell DA, Brenner M, Curtis JH, Guilderson T (2001) Solar forcing of drought frequency in the Maya lowlands. Science 292 1367–1370
Holzhauser H (1997) Fluctuations of the Grosser Aletsch Glacier and the Groner Glacier during the last 3200 years: New results. In Frenzel B, Boulton GS, Glser B, Huckriede U (eds) Glacier fluctuations during the Holocene, Akademie der Wissenschaften und der Literatur, p 35–58
Hong YT, Jiang HB, Liu TS, Zhou LP and 5 others (2000) Response of climate to solar forcing recorded in a 6000-year $\delta^{18}O$ time-series of Chinese peat cellulose. Holocene 10:1–7
Hu FS, Ito E, Brown TA, Curry BB, Engstrom DR (2001) Pronounced climatic variations in Alaska during the last two millenia. Proc Natl Acad Sci USA 98:10552–10556
Huang S, Shen PY, Pollack HN (1996) Deriving century-long trends of surface temperature change from borehole temperatures. Geophys Res Lett 23:257–260
Huang S, Pollack HN, Shen PY (1997) Late Quaternary temperature changes seen in the world-wide continental heat flow measurements. Geophys Res Lett 24:1947–1950
Huang S, Pollack HN, Shen PY (2000) Temperature trends over the past five centuries reconstructed from borehole temperatures. Nature 403:756–758
Huffman TN (1996) Archaeological evidence for climatic change during the last 2000 years in Southern Africa. Quat Int 33:55–60
Hughes MK, Diaz HF (1994) Was there a 'Medieval Warm Period', if so, where and when? Clim Change 26:109–142
Idso SB (1989) Letter to the editor—a problem for paleoclimatology? Quat Res 31:433–434
Ingram MJ, Underhill DJ, Wigley TML (1978) Historical climatology. Nature 276:329–334
Iriondo M (1999) Climatic changes in the South American plains: records of a continent-scale oscillation. Quat Int 57-58:93–112
Jacoby GC, D'Arrigo RD (1995) Tree-ring width and density evidence for climatic and potential forest change in Alaska. Glob Biogeochem Cyc 9:227–234
Jacoby GC, D'Arrigo RD, Davaajamts T (1996) Mongolian tree rings and 20th-century warming. Science 273: 771–773
Jacoby GC, Lovelius NV, Shumilov OI, Raspopov OM, Karbainov JM, Frank DC (2000) Long-term temperature trends and tree growth in the Taymir region of northern Siberia. Quat Res 53:312–318
Jennings AE, Weiner NJ (1996) Environmental change in eastern Greenland during the last 1300 years: evidence from foraminifera and lithofacies in Nansen Fjord, 68° N. Holocene 6:171–191
Jenny B, Valero-Garces B, Urrutia R, Kelts K, Veit H, Appleby PG, Geyh M (2002) Moisture changes and fluctuations of the Westerlies in Mediterranean Central Chile during the last 2000 years: the Laguna Aculeo record (33°–50° S). Quat Int 87:3–18
Johnson TC, Barry S, Chan Y, Wilkinson P (2001) Decadal record of climate variability spanning the past 700 yr in the Southern tropics of East Africa. Geology 29:83–86
Jones PD, Briffa KR, Barnett TP, Tett SFB (1998) High-resolution palaeoclimatic records for the last millennium: interpretation, integration and comparison with General Circulation Model control-run temperatures. Holocene 8: 455–471
Jones PD, New M, Parker DE, Martin S, Rigor IG (1999) Surface air temperatures and its changes over the past 150 years. Rev Geophys 37:173–199
Jones PD, Osborn TJ, Briffa KR (2001) The evolution of climate over the last millennium. Science 292:662–667
Karlén W (1998) Climate variations and the enhanced greenhouse effect. Ambio 27: 270—274
Karlöf L, Winther JG, Isaksson E, Kohler JF and 10 others (2000) A 1500 year record of accumulation at Amundsenisen western Dronning Maud Land, Antarctica, derived from electrical and radioactive measurements on a 120 m ice core. J Geophys Res 105: 12471—12483
Kaser G (1999) A review of the modern fluctuations of tropical glaciers. Glob Planet Change 22:93–103

Keigwin LD (1996) The little ice age and medieval warm period in the Sargasso Sea. Science 274:1504–1508

Kitagawa H, Matsumoto E (1995) Climatic implications of $\delta^{13}C$ variations in a Japanese cedar *Cryptomeria japonica* during the last two millenia. Geophys Res Lett 22:2155–2158

Knapp PA, Soulé PT, Grissino-Mayer HD (2001) Post-drought growth responses of western juniper (*Juniperus occidentalis* var. *occidentalis*) in central Oregon. Geophys Res Lett 28:2657–2660

Knappenberger PC, Michaels PJ, Davis RE (2001) Nature of observed temperature changes across the United States during the 20th century. Clim Res 17:45–53

Koerner RM (1977) Devon Island ice cap: core stratigraphy and paleoclimate. Science 196 15–18

Koerner RM, Fisher DA (1990) A record of Holocene summer climate from a Canadian high-Arctic ice core. Nature 343: 630–631

Korhola A, Weckström J, Holmström L, Erästö P (2000) A quantitative Holocene climatic record from diatoms in Northern Fennoscandia. Quat Res 54:284–294

Kreutz KJ, Mayewski PA, Meeker LD, Twickler MS, Whitlow SI, Pittalwala II (1997) Bipolar changes in atmospheric circulation during the Little Ice Age. Science 277: 1294–1296

Kuhnert H, Pätzold J, Schnetger B, Wefer G (2002) Sea-surface temperature variability in the 16th century at Bermuda inferred from coral records. Palaeogeogr Palaeoclim Palaeoecol 179:159–171

Laird KR, Fritz SC, Maasch KA, Cumming BF (1996) Greater drought intensity and frequency before AD 1200 in the Northern Great Plains, USA. Nature 384:552–554

Lamb HH (1965) The early medieval warm epoch and its sequel. Palaeogeogr Palaeoclim Palaeoecol 1:13–37

Lamb HH (1977) Climate—Past, Present and Future Vol. 2 Climatic History and Future, Methuen

Lamb HH (1982) Climate history and the modern world, Methuen

Lamoureux SF, Bradley RS (1996) A late Holocene varved sediment record of environmental change from northern Ellesmere Island, Canada. J Paleolim 16:239–255

Lara A, Villalba R (1993) A 3260-year temperature record from *Fitzroya cupressoides* tree rings in southern South America. Science 260:1104–1106

Lawton RO, Nair US, Pielke RA Sr, Welch RM (2001) Climatic impact of tropical lowland deforestation on nearby montane cloud forests. Science 294:584–587

Lewis TJ, Wang K (1998) Geothermal evidence for deforestation induced warming: implications for the climatic impact of land development. Geophys Res Lett 25:535–538

Li HC, Bischoff JL, Ku TL, Lund SP, Stot LD (2000) Climate variability in East-Central California during the past 1000 years reflected by high-resolution geochemical and isotopic records from Owens Lake sediments. Quat Res 54: 189–197

Liu K, Yao Z, Thompson LG (1998) A pollen record of Holocene climatic changes from Dunde ice cap, Qinghai-Tibetan Plateau. Geology 26:135–138

Lotter AF, Birks HJB, Eicher V, Hofmann W, Schwander J, Wick L (2000) Younger Dryas Allerød summer temperatures at Gerzensee (Switzerland) inferred from fossil pollen and cladoceran assemblages. Palaeogeogr Palaeoclim Palaeoecol 159:349–361

Luckman BH (2000) The Little Ice Age in the Canadian Rockies. Geomorph 32:357–384

Luckman BH, Briffa KR, Jones PD, Schweingruber FH (1997) Tree-ring based reconstruction of summer temperatures at Columbia Icefield, Alberta, Canada, AD 1073-1983. Holocene 7:375–389

Luterbacher J, Rickli R, Tinguely C, Xoplaki E and 30 others (2000) Monthly mean pressure reconstruction for the late Maunder Minimum period (AD 1675-1715). Int J Clim 20:1049–1066

Madole RF (1994) Stratigraphic evidence of desertification in the west-central Great Plains within the past 1000 yr. Geology 22:483–486

Manley G (1974) Central England temperatures: monthly means 1659 to 1973. Quart J Roy Meteorol Soc 100: 389–405

Mann ME, Bradley RS, Hughes MK (1998) Global-scale temperature patterns and climate forcing over the past six centuries. Nature 392:779–787

Mann ME, Bradley RS, Hughes MK (1999) Northern Hemisphere temperatures during the past millennium: inferences, uncertainties, and limitations. Geophys Res Lett 26:759–762

Mann ME, Grille E, Bradley RS, Hughes MK, Overpeck J, Keimig FT, Gross W (2000) Global temperature patterns in past centuries: an interactive presentaion. Earth Interact 4:1–29 [available online http://Earthinteraction.org]

Marshall GJ, Lagun V, Lachlan-Cope TA (2002) Changes in Antarctic Peninsula tropospheric temperatures from 1956 to 1999: a synthesis of observations and reanalysis data. Int J Clim 22:291–310

Marshall J, Johnson H, Goodman J (2001) A study of the interaction of the North Atlantic Oscillation with ocean circulation. J Clim 14:1399–1421

McDermott F, Mattey DP, Hawkesworth C (2001) Centennial-scale Holocene climate variability revealed by a high-resolution speleothem $\delta^{18}O$ record from SW Ireland. Science 294:1328–1331

Meese DA, Gow AJ, Grootes P, Mayewski PA and 5 others (1994) The accumulation record from the GISP2 core as an indicator of the climate change throughout the Holocene. Science 266: 1680—1682

Mercer JH (1978) West Antarctic ice sheet and CO_2 greenhouse effect: a threat of disaster. Nature 271:321–325

Morgan VI (1985) An oxygen isotope climate record from the Law Dome, Antarctica. Clim Change 7:415–426

Mosley-Thompson E (1995) Paleoenvironmental conditions in Antarctica since A.D. 1500: ice core evidence. In: Bradley RS, Jones PD (eds) Climate since A.D. 1500, Routledge, New York, p 572–591

Mosley-Thompson E, Thompson LG (1982) Nine centuries of microparticles deposition at the South Pole. Quat Res 17:1–13

Naurzbaev MM, Vaganov EA (2000) Variation of early summer and annual temperature in east Taymir and Putorian (Siberia) over the last two millenia. J Geophys Res 105: 7317—7326

Ogilvie AEJ (1984) The past climate and sea-ice record from Iceland. Part 1. Data to A.D. 1780. Clim Change 6:131–152

Ogilvie A, Farmer G (1997) Documenting the medieval climate. In: Hulme M, Barrow E (eds) Climate of the British Isles—present, past and future. Routledge, London, p 112–133

Ogilvie AEJ, Jónsson T (2001) 'Little Ice Age' research: a perspective from Iceland. Clim Change 48:9–52

Ogilvie AEJ, Barlow LK, Jennings AE (2000) North Atlantic climate c. A.D. 1000: millennial reflections on the Viking discoveries of Iceland, Greenland and North America. Weather 55:34–45

Overpeck J, Hughen K, Hardy D, Bradley R 14 others (1997) Arctic environmental change of the last four centuries. Science 278:1251–1256

Pederson JL (2000) Holocene paleolakes of Lake Canyon,

Colorado Plateau: Paleoclimate and landscape response from sedimentology and allostratigraphy. Geol Soc Amer Bull 112:147–158
Perry CA, Hsu KJ (2000) Geophysical, archaeological, and historical evidence support a solar-output model for climate change. Proc Natl Acad Sci USA 97:12433–12438
Petersen KL (1994) A warm and wet little climatic optimum and a cold and dry little ice age in the Southern Rocky Mountains, USA. Clim Change 26:243–269
Pfister C (1995) Monthly temperature and precipitation in central Europe 1525–1979: quantifying documentary evidence on weather and its effects. In: Bradley RS, Jones PD (eds) Climate since A.D. 1500. Routledge, New York, p 118–142
Pfister C, Luterbacher J, Schwarz-Zanetti G, Wegmann M (1998) Winter air temperature variations in western Europe during the early and high Middle ages (AD 750–1300). Holocene 8:535–552
Pisias NG (1978) Paleoceanography of the Santa Barbara Basin during the last 8000 years. Quat Res 10:366–384
Porter SC (1981) Glaciological evidence of holocene climatic change. In: Wigley TML, Ingram MJ, Farmer G (eds) Climate and history—studies in past climates and their impact on man. Cambridge University Press, Cambridge, p 82–110
Porter SC (1986) Pattern and forcing of the Northern Hemisphere glacier variations during the last millennium. Quat Res 26:27–48
Proctor CJ, Baker A, Barnes WL, Gilmour MA (2000) A thousand year speleothem proxy record of North Atlantic climate from Scotland. Clim Dyn 16:815–820
Ramesh R (1993) First evidence for Little Ice Age and Medieval Warming in India (abstract). EOS Trans AGU suppl 74: 118
Ren G (1998) Pollen evidence for increased summer rainfall in the Medieval Warm Period at Maili, Northeast China. Geophys Res Lett 25:1931–1934
Robinson WA, Hansen J, Ruedy R (2001) Where's the heat? Insights from GCM experiments into the lack of Eastern U.S. warming. American Meteor Soc 81st Annual Meeting, Albuquerque NM, January 14–19 2001 (abstract) American Meteorological Society, Washington DC
Rodrigo FS, Esteban-Parra MJ, Pozo-Vázquez D, Castro-Díez Y (2000) Rainfall variability in southern Spain on decadal to centennial time scales. Int J Clim 20:721–732
Rott H, Skvarca P, Nagler T (1996) Rapid collapse of northern Larsen ice shelf, Antarctica. Science 271:788–792
Schleser GH, Helle G, Lücke A, Vos H (1999) Isotope signals as climate proxies: the role of transfer functions in the study of terrestrial archives. Quat Sci Rev 18:927–943
Schmutz C, Luterbacher J, Gyalistras D, Xoplaki E, Wanner H (2000) Can we trust proxy-based NAO index reconstructions? Geophys Res Lett 27:1135–1138
Schuster PF, White DE, Naftz DL, Cecil LD (2000) Chronological refinement of an ice core record at Upper Fremont Glacier in south central North America. J Geophys Res 105:7317–7326
Scuderi LA (1993) A 2000-year tree ring record of annual temperatures in the Sierra Nevada mountains. Science 259: 1433–1436
Seager R, Kushnir Y, Visbeck M, Naik N, Miller J, Krahmann G, Cullen H (2000) Causes of Atlantic ocean climate variability between 1958 and 1998. J Clim 13:2845–2862
Serre-Bachet F (1994) Middle ages temperature reconstructions in Europe, a focus on Northeastern Italy. Clim Change 26:213–240
Shi Y, Yao T, Yang B (1999) Decadal climatic variations recorded in Guliya ice core and comparison with the historical documentary data from East China during the last 2000 years. Sci in China Ser D (suppl) 42:91–100
Skinner WR, Majorowicz JA (1999) Regional climatic warming and associated twentieth century land-cover changes in north-western North America. Clim Res 12:39–52
Slonosky VC, Yiou P (2001) The North Atlantic Oscillation and its relationship with near surface temperature. Geophys Res Lett 28:807–810
Smith RC, Ainley D, Baker K, Domack E and 7 others (1999) Marine ecosystem sensitivity to climate change. Bioscience 49:393–404
Song J (2000) Changes in dryness/wetness in China during the last 529 years. Int J Clim 20:1003–1015
Stahle DW, Cleaveland MK, Hehr JG (1988) North Carolina climate changes reconstructed from tree rings: A.D. 372 to 1985. Science 240: 1517–1519
Stine S (1994) Extreme and persistent drought in California and Patagonia during medieval time. Nature 369:546–549
Stine S (1998) Medieval climatic anomaly in the Americas. In: Issar AS, Brown N (eds) Water, environment and society in times of climatic change. Kluwer Academic Press, Dordrecht, p 43–67
Stuiver M, Grootes PM, Braziunas TF (1995) The GISP2 $\delta^{18}O$ climate record of the past 16,500 years and the role of the Sun, ocean and volcanoes. Quat Res 44:341–354
Swetnam TW (1993) Fire history and climate change in Giant Sequoia groves. Science 262:885–889
Tagami Y (1993) Climate change reconstructed from historical data in Japan. In: Proc Int Symp Global Change by IGBP March 27–29 1992, Tokyo. International Geosphere-Biosphere Programme-IGBP, Tokyo, p 720–729
Tagami Y (1996) Some remarks on the climate in the Medieval Warm Period of Japan. In: Mikami T, Matsumoto E, Ohta S, Sweda T (eds) Paleoclimate and environmental variability in Austral-Asian transect during the past 2000 years (IGBP Proceedings). International Geosphere-Biosphere Programme-IGBP, Nagoya, p 115–119
Tarussov A (1995) The Arctic from Svalbard to Severnaya Zemlya: climatic reconstructions from ice cores. In Bradley RS, Jones PD (eds) Climate Since A.D. 1500, Routledge, New York, p 505–516
Thomas RH, Sanderson TJO, Rose KE (1979) Effect of climatic warming on the West Antarctic ice sheet. Nature 277: 355–358
Thompson LG (2000) Ice core evidence for climate change in the Tropics: implications for our future. Quat Sci Rev 19: 19–35
Thompson LG, Mosley-Thompson E, Dansgaard W, Grootes PM (1986) The little ice age as recorded in the stratigraphy of the tropical Quelccaya ice cap. Science 234:361–364
Thompson LG, Mosley-Thompson E, Davis ME, Bolzan JF and 6 others (1989) Holocene-Late Pleistocene climatic ice core records from Qinghai-Tibetan plateau. Science 246:474–477
Thompson LG, Peel DA, Mosley-Thompson E, Mulvaney R, Dai J, Lin PN, Davis ME, Raymond CF (1994) Climate since AD 1510 on Dyer Plateau, Antarctic Peninsula: evidence for recent climate change. Ann Glaciol 20:420–426
Thompson LG, Mosley-Thompson E, Davis ME, Lin PN, Dai J, Bolzan JF, Yao T (1995) A 1000 year climate ice-core record from the Guliya ice cap, China: its relationship to global climate variability. Ann Glaciol 21:175–181
Thompson LG, Yao T, Mosley-Thompson E, Davis ME, Henderson KA, Lin PN (2000) A high-resolution millennial record of the South Asian Monsoon from Himalayan ice cores. Science 289:1916–1919

Till C, Guiot J (1990) Reconstruction of precipitation in Morocco since 1100 A.D. based on *Cedrus atlantica* tree-ring widths. Quat Res 33:337–351

Tong G, Zhang J, Fan S, Zhao J, Wang G, Wang Y, Wang L, Cui J (1996) Environmental changes at the top of Taibai Mountain, Qinling since 1ka ago. Mar Geol Quat Geol 16: 95—104

Tyson PD, Lindesay JA (1992) The climate of the last 2000 years in southern Africa. Holocene 2:271–278

Tyson PD, Karlén W, Holmgren K, Heiss GA (2000) The Little Ice Age and medieval warming in South Africa. South African J Sci 96:121–126

Valero-Garcés BL, Delgado-Huertas A, Ratto N, Navas A, Edwards L (2000) Paleohydrology of Andean saline lakes from sedimentological and isotopic records, Northwestern Argentina. J Paleolimnology 24:343–359

van Engelen AFV, Buisman J, Ijnsen F (2001) A millennium of weather, winds and water in the low countries. In: Jones PD, Ogilvie AEJ, Davies TD, Briffa KR (eds) History and climate—memories of the future? Kluwer Academic Publishers, New York, p 101–124

Vaughan DG, Doake CSM (1996) Recent atmospheric warming and retreat of ice shelves on the Antarctic Peninsula. Nature 379:328–331

Verardo S, Cronin T, Willard D, Kerhin R (1998) Climate and environmental changes recorded by Chesapeake Bay Dinoflagellate cysts in the last millennium. Geol Soc Amer Abs 74: 284

Verschuren D, Laird KR, Cumming BF (2000) Rainfall and drought in equatorial east Africa during the past 1,100 years. Nature 403:410–414

Editorial responsibility: Chris de Freitas, Auckland, New Zealand

Villalba R (1990) Climatic fluctuations in northern Patagonia during the last 1000 years as inferred from tree-ring records. Quat Res 34:346–360

Villalba R (1994) Tree-ring and glacial evidence for the medieval warm epoch and the little ice age in southern South America. Clim Change 26:183–197

von Storch JS, Müller P, Bauer E (2001) Climate variability in millennium integrations with coupled atmosphere-ocean GCMs: a spectral view. Clim Dyn 17:375–389

Williams PW, Marshall A, Ford DC, Jenkinson AV (1999) Palaeoclimatic interpretation of stable isotope data from holocene speleothems of the Waitomo district, North Island, New Zealand. Holocene 9:649–657

Wilson AT, Hendy CH, Reynolds CP (1979) Short-term climate change and New Zealand temperatures during the last millennium. Nature 279:315–317

Winkler S (2000) The 'Little Ice Age' maximum in the southern Alps, New Zealand: preliminary results at Mueller Glacier. Holocene 10:643–647

Winter A, Ishioroshi H, Watanabe T, Oba T, Christy J (2000) Carribean sea surface temperatures: two-to-three degrees cooler than present during the Little Ice Age. Geophys Res Lett 27:3365–3368

Woodhouse CA, Overpeck JT (1998) 2000 years of drought variability in the central United States. Bull Am Meteor Soc 79:2693–2714

Yu Z, Ito E (1999) Possible solar forcing of century-scale drought frequency in the northern Great Plains. Geology 27:262–266

Zhang D (1994) Evidence for the existence of the medieval warm period in China. Clim Change 26:289–297

Submitted: April 11, 2002; Accepted: August 29, 2002
Proofs received from author(s): January 3, 2003

GEOPHYSICAL RESEARCH LETTERS, VOL. 31, L03209, doi:10.1029/2003GL019141, 2004

Estimation and representation of long-term (>40 year) trends of Northern-Hemisphere-gridded surface temperature: A note of caution

Willie W.-H. Soon,[1] David R. Legates,[2] and Sallie L. Baliunas[1]

Received 24 November 2003; revised 17 December 2003; accepted 24 December 2003; published 14 February 2004.

[1] Several quantitative estimates of surface instrumental temperature trends in the late 20th century are compared by using published results and our independent analyses. These estimates highlight a significant sensitivity to the method of analysis, the treatment of data, and the choice of data presentation (i.e., size of the smoothing filter window). Providing an accurate description of both quantitative uncertainties and sensitivity to the treatment of data is recommended as well as avoiding subjective data-padding procedures. *INDEX TERMS:* 1620 Global Change: Climate dynamics (3309); 3299 Mathematical Geophysics: General or miscellaneous; 1699 Global Change: General or miscellaneous; 6620 Public Issues: Science policy. **Citation:** Soon, W. W.-H., D. R. Legates, and S. L. Baliunas (2004), Estimation and representation of long-term (>40 year) trends of Northern-Hemisphere-gridded surface temperature: A note of caution, *Geophys. Res. Lett.*, *31*, L03209, doi:10.1029/2003GL019141.

1. Introduction

[2] It has been estimated that the Earth's composite air and sea surface temperature increased by approximately 0.6 ± 0.2°C in the 20th century [*IPCC TAR*, 2001, Third Assessment Report's (TAR) chapter 2]. These same analyses also suggest that the surface warming may have accelerated during the last two decades of the 20th century, with a linear 20-year trend of about 0.10 to 0.20 ± 0.07°C per decade, depending on the particular data-set adopted (as summarized in Table 2.1, 2.2 and 2.3 of *IPCC TAR*, 2001). While the error bars represent a plausible assessment of "uncertainties in the annual anomalies due to data gaps, urbanisation over land, and bias correction to SST [sea surface temperature]" [*IPCC TAR*, 2001, 115, which refers to work by *Folland et al.*, 2001], more robust and reliable methods for the quantification of sampling biases, errors and uncertainties for the construction of local, regional and global-scale climatologies are needed [e.g., *Christy et al.*, 2001; *Pielke et al.*, 2002; *Arnfield*, 2003]. Analyses such as *Folland et al.* [2001] that rely on climate model results should be avoided because current climate models are deficient in representing the current climate [e.g., *Lindzen*, 1994; *Johnson*, 1997; *Pielke*, 2001].

[3] We appreciate, and here leave aside, the unresolved questions concerning the many recent efforts to relate the air temperature variabilities recorded by these short-term instrumental records to the broader perspective of historical changes in the climate system over the past millennium or so through natural paleoclimatic indicators [see for example exchanges between *Mann et al.*, 2003 and *Soon et al.*, 2003]. But issues regarding the resolution, detection and quantification of long-term (i.e., multidecadal or centennial) temperature trends, even through the use of instrumental records, are far from settled. Figure 1 illustrates this point with a collection of published figures by (a) *IPCC TAR* [2001], (b) *Mann* [2002], (c) *Mann et al.* [2003] and (d) *Mann and Jones* [2003]. Figure 1 highlights both the remarkably warm level and the rapid rate of warming of the last two decades of the 20th century that have been presented in the smoothed or filtered instrumental record of the Northern Hemisphere. Further, visual inspection shows a sudden acceleration to about 0.25°C per year in the period 2002 through 2003 between the lowest estimate [*Mann*, 2002] and the highest [*Mann and Jones*, 2003].

[4] Does the accelerated warming trend presented in Figure 1d simply depend on the smoothing or filtering process, which relies sensitively on the details of the application?

[5] The objectives of this paper are:

[6] (1) to inter-compare several general methods of representing instantaneous long-term trends estimated from the same temperature record but yielding significantly different values, as assembled in Figure 1. Trends are defined to involve time scales greater than 40 years (to conform to the results in Figure 1) and trends have been estimated from (i) a simple running mean, (ii) a smoothing with the use of a more sophisticated filter such as the Hamming filter as prescribed for the results in IPCC TAR (Figure 1a) and (iii) a frequency-filtering technique, namely, the wavelet transform; and

[7] (2) to ask if one can emulate the instrumental data trends shown in Figure 1, and if so under what assumptions, and if not, why not.

2. Data and Methods

[8] The data used here are the combined terrestrial air and sea surface temperature record for the Northern Hemisphere (downloaded November 4, 2003 from the public access URL of the Climate Research Unit at the University of East Anglia: http://www.cru.uea.ac.uk/cru/data/temperature/.). Only the annual-mean anomaly (relative to the reference period 1961–1990) series from 1856 through 2002 were used for our study concerning estimates of long-term (>40 year) warming or cooling trends. The annual temperature data are accompanied by error bars of two standard deviations (±0.05°C) for the period after 1951 with the formal uncertainties expanded to four times larger during the 1850s.

[1]Harvard-Smithsonian Center for Astrophysics, Cambridge, Massachusetts, USA.
[2]Center for Climatic Research, University of Delaware, Newark, Delaware, USA.

0094-8276/04/2003GL019141$05.00

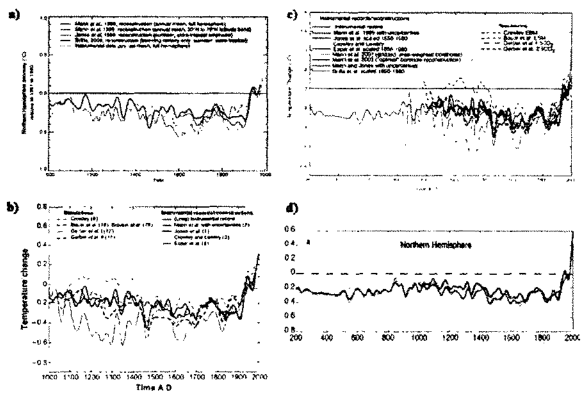


Figure 1. Published Northern Hemispheric surface temperature trends that suggest an unprecedented (relative to the past 1000 to 1800 years) warming of the late 20th century as shown by the smoothed or filtered thermometer records in (a) *IPCC TAR* [2001, Chapter 2, p. 134], (b) *Mann* [2002], (c) *Mann et al.* [2003] and (d) *Mann and Jones* [2003]. Note both the similarity of the smoothed instrumental series in all panels for the mid-to-late 19th century to early 20th century interval but a significant difference in the actual heights of the trend line at 2000, especially between (b) and (d). (Insert (a) is reprinted with permission from the IPCC; Insert (b) is reprinted with permission from *Mann* [2002], Copyright 2002, AAAS (http://www.sciencemag.org); Inserts (c) and (d) are reproduced with permission from American Geophysical Union).

[9] The three methods of trend estimation and representation adopted here are standard approaches [e.g., *Robeson*, 1997], with the exception of the frequency-filtering technique. This last method uses a version of the adaptive wavelet transform algorithm by *Frick et al.* [1998] that has been previously applied to characterize and quantify the variability in Central England air temperature records [*Baliunas et al.*, 1997]. The Hamming-weight filter [$w(i) = 0.54 - 0.46\cos\{2\pi i/(n-1)\}$ where i is the year/data point within the window and n is the length of the window] was selected by the *IPCC TAR* (see their Figure 2.21, p. 134): "All series were smoothed with a 40-year Hamming-weights lowpass filter, with boundary constraints imposed by padding the series with its mean values during the first and last 25 years." We note that alternative filter window widths of 10, 20, 30 or even 5 and 50 years had all been applied by various authors to characterize the long-term Northern Hemisphere temperature trend in recent publications. This introduces subjectivity in the quantification of long-term air temperature trends. We shall adopt a 40-year window width for a more direct comparison with published trends of Figure 1.

[10] In this short paper, we do not discuss the currently unresolved issues concerning the different air temperature trends deduced from surface thermometers and from satellite-based and/or balloon-based measurements, especially over the last 20–25 years. We also do not discuss the physical arguments necessary for justification in considering long-term temperature trends, rather than seasonal and interannual or even decadal-scale changes, for the association with persistent climatic forcings like the enhanced level of atmospheric carbon dioxide.

3. Results and Discussion

[11] Figure 2 shows results from the three different methods of estimating and representing long-term (>40 year) trends in the gridded annual surface temperature of the Northern Hemisphere from 1856–2002. The simple running mean (Figure 2a) and wavelet-transform frequency filtering (Figure 2c) cases were performed with no data padding, while the Hamming-filter weighting (Figure 2b) case prescribed data padding with mean values of the first and last 25 years of the record at left and right end points, respectively (red curve). The last procedure traces the steps taken by IPCC author(s) for producing Figure 2.21 in TAR (p. 134), shown here as Figure 1a.

[12] We successfully replicated the TAR's early 20th century curve presented (Figure 1a,) but not its late 20th century trend, even when the exact time interval specified by the TAR, 1900–1999, was used. The last point of the 20th century is too low in our calculation (0.25°C if endpoints were filled after 1999 [not plotted here] or >0.3°C if endpoints were unjustifiably filled after 2002, shown in Figure 2b) while the *IPCC TAR* Figure 2.21 (Figure 1a here) reaches a significantly higher value, of at least 0.4°C, by 1999. We were able to get the high value of >0.4°C (green curve in Figure 2b) if we padded the forward endpoint with the value of 2002 temperature anomaly (thus we also padded the backward end point with the 1856

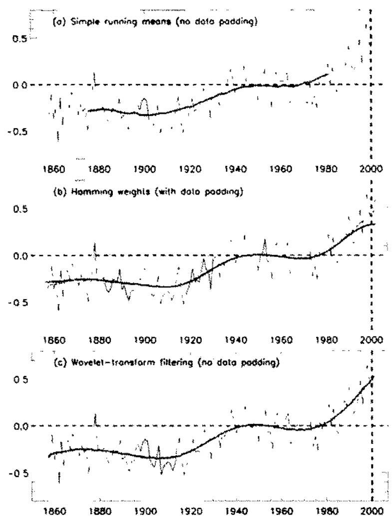


Figure 2. The yearly mean Northern Hemisphere surface temperature anomaly from 1856 through 2002 (blue) superposed with trend estimates (red) using (a) a simple 40-yr running mean with no padding of endpoints; (b) a 40-yr Hamming-weight smoothing with padding of endpoints with mean values of first and last 25 years so that the trend line can be represented at both endpoints of the original series (at 1856 and 2002) as discussed in the caption of Figure 2.21 of IPCC TAR; and (c) a filtered series removing all timescales between 2 and 40 years using the wavelet transform. The additional trend estimate (green curve) in (b) represents the alternative scenario of padding the respective endpoints with values at 1856 and 2002 rather than the 25-year means specified by IPCC TAR. The horizontal zero-line (dashed curve) represents the mean anomaly of the instrumental base period 1961–1990. Note that significant quantitative differences exist in the various representations of the trend curves near the year 2000. This results from different data smoothing methods and data handlings used near the forward endpoint of the curve.

anomaly for 20 years, which makes a slight change in the curve), even though the 2002 data were unavailable when TAR was prepared. But this disagreement with TAR neither implies nor proves that our calculations are wrong. As noted below, we essentially recover, by Hamming-filter weighting, the curves in *Mann* [2002] (as Figure 1b here and red curve in Figure 2b) and *Mann* [2000] (not shown here). The discrepancies seen in Figure 1a and Figure 2b cannot be explained by an argument that the *IPCC TAR* author(s) used a different data set because we can recover almost exactly the near-decadally-smoothed curve in *IPCC TAR*'s Figure 2.7a on p. 114 (not shown here). Perhaps TAR performed other analyses that were not discussed.

[13] The difference in the estimated trend of the late 20th century between Figure 1d and Figure 1b, about 0.25°C per year, is very large and if physically real would have an important implication for global warming discussions. The trend by 2000 of >0.4°C published by *Mann et al.* [2003] (our Figure 1c) is similar to the TAR result (our Figure 1a), but significantly below the >0.5°C value in *Mann and Jones* [2003; our Figure 1d].

[14] We were unable to reproduce the smoothed Northern Hemisphere temperature trend of *Mann and Jones* [2003, their Figure 2a or Figure 1d here], which is puzzling because we did recreate their smoothed Southern Hemisphere temperature trend and its amplitude (their Figure 2b not shown here). We were able to do so by applying the 40-year Hamming-weighted filter to the University of East Anglia's Southern Hemisphere land and sea surface temperature record from 1856–2002, not as specified by TAR, but by padding the respective endpoints backward with the value at 1856 and forward with value at 2002.

[15] Figures 3a, 3b, and 3c show the progression toward increasingly higher values at 2000 over the short 1–2 years spans from the publications of Figure 1b to Figures 1c and 1d. Figure 3d illustrates the same quantities as in Figure 2b and Figure 2c, but now with both axes adjusted to more closely resemble the most recent 200 years of Figure 1d. Our wavelet-transform lowpass filtered curve (red) resembles, though does not exactly correspond to, Figure 1d [*Mann and Jones*, 2003] but there is no indication that Figure 1d was derived from such method of trend estimation. Our 40-year Hamming-weight smoothed curve

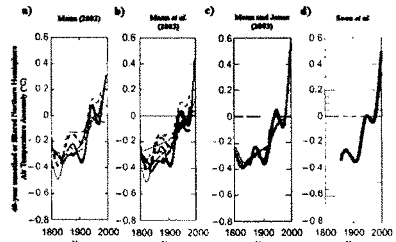


Figure 3. The comparison of the 40-yr low-pass filtered series using wavelet transform (red curve, same as Figure 2c) with the 40-yr Hamming weights smoothed series (blue curve, same as Figure 2b) with endpoints padding recommended by the author(s) of Figure 2.21 of the IPCC TAR. Our curves (panel d) are, in turn, compared on the same time-temperature scales with smoothed series (all the red-solid curves) by (a) *Mann* [2002] (Figure 1b), (b) *Mann et al.* [2003] (Figure 1c), and (c) *Mann and Jones* [2003] (Figure 1d). A progression is seen toward increasingly higher values at 2000 over the short 1–2 years span of publication interval. Figure (3a) is reprinted with permission from *Mann* [2002], Copyright 2002, AAAS (http://www.sciencemag.org).

(with padding of endpoints with respective mean values of first and last 25 years of data record) is similar to Figure 1b [*Mann*, 2002] (and results in *Mann*, 2000) and this is likely the method of smoothing performed in these articles.

[16] Quantification of the late 20th-century Northern Hemisphere temperature trend is highly sensitive to the choice of methodology and treatment of data. Padding additional data at the endpoints (as performed in Figure 2b and TAR) simply to fill in a trend line to cover the full extent of the actual series (say, 1856 through 2002) is unphysical and can be misleading. The data-padding procedure is the logical equivalent of perfect knowledge in retrodicting the past and predicting the future.

[17] The remaining methods (Figure 2a and 2c) give two extreme results. The running mean without data padding (Figure 2a) is agnostic about the present, and argues that the actual trend of the late 20th century (smoothed) variability will be decided by what occurs in the coming decades. The running mean does offer a very consistent framework for the objective standard for trend estimation and representation, even though the most recent data may visually suggest a different trend.

[18] Although the wavelet-transform frequency-filtering method (Figure 2c) is the least sensitive to treatment of endpoints (yet is quite sensitive in some realistic scenarios we examined offline), temporally local information may be overemphasized, hence biasing the long-term trend value. Thus, there is no compelling physical argument to prefer the wavelet representation to the running mean (Figure 2a).

[19] *All* series in Figure 1 have similar, if not exact, instrumental trends and data values during the mid-to-late 19th century to early 20th century, while trend differences in Figure 1 point to significant differences in the imposed endpoint conditions of the late 20th century. We note that too little time has passed between the publication of the several articles to expect significant changes to the input temperature data. A close study of Figure 2b and 2c suggests the difference of trends at the end of the 20th century (i.e., about 0.25°C in Figure 1 [between Figure 1d and Figure 1b] and about 0.2°C in Figure 2 [between red curves in Figure 2b and 2c]) can be explained by the method of trend representation and/or the treatment of data at the forward endpoint (i.e., the updating of record with new data points) before smoothing or filtering rather than any real physical changes. Additional sensitivity tests (not shown here) support this explanation. We find no justification for the apparently spurious, high value of the late 20th-century trend in Figure 1d.

4. Conclusions

[20] We were successful in replicating the instrumental surface temperature trends in several previously published results. However, we failed to reproduce the long-term (>40 years) Northern Hemispheric surface thermometer temperature trends shown in Figure 2.21 of *IPCC TAR* (as Figure 1a here) and in Figure 2a of *Mann and Jones* [2003] (as Figure 1d here). We conclude that published results suggesting that the Northern Hemisphere surface air temperature has increased by the extremely rapid rate of about 1 to 2.5°C per decade during the last one year (2002–2003) (see Figure 3) are most likely artefacts of methodology and procedure of trend smoothing. Accurate communication of methods and avoidance of data-padding procedures for smoothing and/or filtering of climatic time series should be incorporated in reporting data trends.

[21] **Acknowledgments.** We thank Gene Avrett, Steve Cranmer and Daniel Leathers for comments that helped improve the paper. We thank our colleague Peter Frick of the Institute of Continuous Media Mechanics, Perm, Russia for his sharing of expertise in the development of the gapped (or data adaptive) wavelet transform algorithm. This work was partially supported by the Air Force Office of Scientific Research Grant AF49620-02-1-0194.

References

Arnfield, A. J. (2003), Two decades of urban climate research: A review of turbulence, exchanges of energy and water, and the urban heat island, *Int. J. Climatology*, *23*, 1–26.

Baliunas, S., P. Frick, D. Sokoloff, and W. Soon (1997), Time scales and trends in the Central England Temperature data (1659–1990), A wavelet analysis, *Geophys. Res. Lett.*, *24*(11), 1351–1354.

Christy, J. R., D. E. Parker, S. J. Brown, I. Macadam, M. Stendel, and W. B. Norris (2001), Differential trends in tropical sea surface and atmospheric temperatures since 1979, *Geophys. Res. Lett.*, *28*(1), 183–186.

Folland, C. K., N. A. Rayner, S. J. Brown, T. M. Smith, S. S. P. Shen, D. E. Parker, I. Macadam, P. D. Jones, R. N. Jones, N. Nicholls, and D. M. H. Sexton (2001), Global temperature change and its uncertainties since 1861, *Geophys. Res. Lett.*, *28*(13), 2621–2624.

Frick, P., A. Grossmann, and P. Tchamitchian (1998), Wavelet analysis of signals with gaps, *J. Math. Phys.*, *39*, 4091–4107.

IPCC TAR (Intergovernmental Panel on Climate Change, Third Assessment Report) (2001), *Climate Change 2001: The Scientific Basis*, J. T. Houghton, Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, X. Dai, K. Maskell, and C. A. Johnson (Eds.), Cambridge Univ. Press.

Johnson, D. R. (1997), 'General coldness of climate models' and Second Law: Implications for modeling the Earth system, *J. Climate*, *10*, 2826–2846.

Lindzen, R. S. (1994), Climate dynamics and global change, *Annual Rev. Fluids Mech.*, *26*, 353–378.

Mann, M. E. (2000), Lessons for a new millennium, *Science*, *289*, 253–254.

Mann, M. E. (2002), The value of multiple proxies, *Science*, *297*, 1481–1482.

Mann, M. E., and P. D. Jones (2003), Global surface temperatures over the past two millennia, *Geophys. Res. Lett.*, *30*(15), 1820, doi:10.1029/2003GL017814.

Mann, M. E., et al. (2003), On Past Temperatures and Anomalous Late-20th Century Warmth, *Eos*, *84*(27), 256–257.

Pielke, R. A., Sr. (2001), Earth system modeling—An integrated assessment tool for environmental studies, in *Present and future of modeling global environmental change: Toward integrated modeling*, edited by T. Matsuno and H. Kida, TERRAPUB, 311–337, Tokyo.

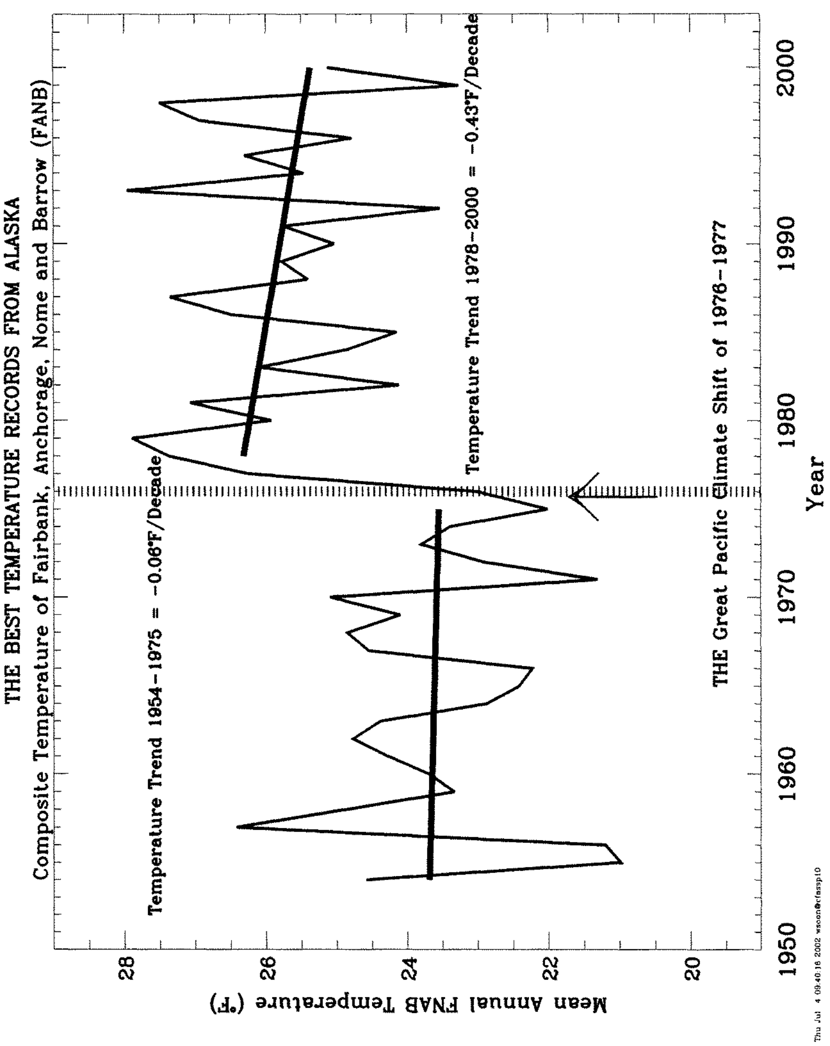
Pielke, R. A., Sr., T. Stohlgren, L. Schell, W. Parton, N. Doesken, K. Redmond, J. Moeny, and T. G. F. Kittel (2002), Problems in evaluating regional and local trends in temperature: An example from eastern Colorado, USA, *Int. J. Climatol.*, *22*(4), 421–434.

Robeson, S. (1997), Statistical Considerations, in *Applied Climatology: Principles and Practices*, edited by R. D. Thompson and A. Perry, Routledge, 22–35, London, UK.

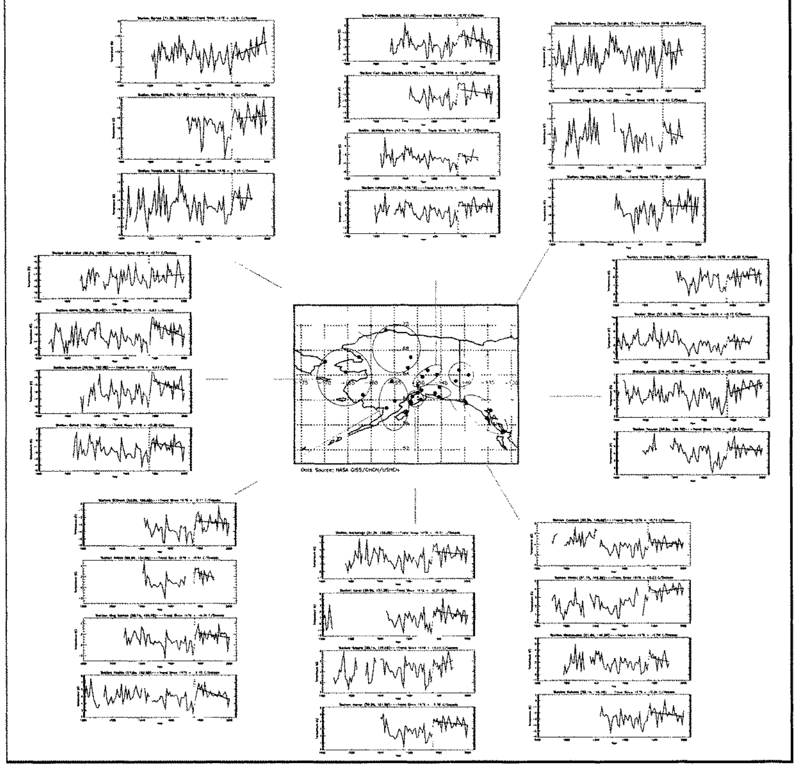
Soon, W., S. Baliunas, and D. R. Legates (2003), Comment on "On Past Temperatures and Anomalous Late-20th Century Warmth," *Eos*, *84*(44), 473–476.

W. W.-H. Soon and S. L. Baliunas, Harvard-Smithsonian Center for Astrophysics, Cambridge, MA, USA. (wsoon@cfa.harvard.edu; sbaliunas@cfa.harvard.edu)

D. R. Legates, Center for Climatic Research, University of Delaware, Newark, DE, USA. (legates@udel.edu)


THE BEST TEMPERATURE RECORDS FROM ALASKA
Composite Temperature of Fairbank, Anchorage, Nome and Barrow (FANB)
Temperature Trend 1954-1975 = -0.06°F/Decade
Temperature Trend 1978-2000 = -0.43°F/Decade
THE Great Pacific Climate Shift of 1976-1977
Mean Annual FNAB Temperature (°F)
28
26
24
22
20
1950
1960
1970
1980
1990
2000
Year

Alaska's Temperature History



RESPONSE BY DR. WILLIE SOON TO ADDITIONAL QUESTIONS FROM SENATOR JEFFORDS

Question 1. In testimony, you said that you did not know whether you submitted something for publication to Capitalism magazine. Here is the title and web address: "Global Warming Speculation vs. Science: Just Ask the Experts" by Sallie Baliunas & Willie Soon (Capitalism Magazine—August 22, 2002) *http://capmag.com/article.asp?ID=1816.* Did you submit or approve submission of this article for publication?

Response. With the benefit of your reminder, I hereby confirm that the above mentioned article in Capitalism Magazine was taken from the original article "Just Ask the Experts" by Baliunas and Soon originally published by the TechCentralStation.com at the link: *http://www.techcentralstation.com/072302B.html.* I did not submit the article to Capitalism Magazine.

Question 2. In your testimony you indicated that your training is in "atmospherics." Could you please explain this term more fully, and indicate your formal training in paleoclimatic studies and analysis?

Response. My PhD thesis[1] was on collisional-radiative properties of high-temperature, partially ionized nitrogen, oxygen, helium and hydrogen plasmas at conditions relevant to the Earth's atmosphere. This is why I mentioned that I had formal training in "atmospheric and space physics" in my oral remarks. If necessary, please consult my thesis advisor, Professor Joseph Kunc at *kunc@usc.edu* for further details about my educational background.

I would add that the quality of knowledge about climate science or any other subject of interest must be judged on its own merits, and does not and must not be determined by invoking the amount of formal schooling or consensus viewpoints adopted by particular interest groups.

My research interests and learning about paleoclimatology has been obtained mainly through the following individuals and sources:

(1) Professor Eric Posmentier (*Eric.S.Posmentier@Dartmouth.EDU*), who is also my colleague.

(2) Professor David Legates (*legates@UDel.Edu*), who is also my colleague.

(3) Participation, both as a student and as lecturer, in numerous national and international workshops, conferences and summer schools including (a) the 1993's NATO Advanced Research workshop on "Solar engine and its influence on terrestrial atmosphere and climate", (b) the 1994's NASA-NOAA Summer School on Processes of Global Change, (c) the 1996's (French) CNRS "Chaos et Fractales dans l'activite Solaire", (d) the 2000's "1st Solar and Space Weather Euroconference: The Solar Cycle and Terrestrial Climate," and other specialized meetings.

(4) Many other scientists also have been helpful in my eager learning of the subject: the late Professor Jean Grove (Girton College, Cambridge University), Professor Jim Kennett (University of California Santa Barbara), Professor David J. A. Evans (University of Glasgow), Professor Lowell Stott (University of Southern California), Professor Hong-Chun Li (University of Southern California), Professor Reid Bryson (University of Wisconsin), Professor Henri Grissino-Mayer (University of Kentucky), Professor Emi Ito (University of Minnesota), Dr. ShaoPeng Huang (University of Michigan), Dr. Zhonghui Liu (Brown University), Dr. Ming Tang (Institute of Geology and Geophysics, Chinese Academy of Sciences), Dr. Yang Bao (Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences), and Professor Bin Wang (University of Hawaii).

Question 3. Do you maintain that the proxy-based temperature reconstructions of the Mann and colleagues do not extend into the latter half of the 20th century?

Response. The proxy-based temperature reconstructions for the Northern Hemisphere by Mann *et al.* (1998, Nature, vol. 392, 779–782) and Mann *et al.* (1999, *Geophysical Research Letters*, vol. 26, 759–762) extend from 1400–1980 and 1000–1980, respectively. So it is true that those proxy-based temperature series did not cover the 1981–2000 interval of the late 20th century.

Here is what close colleagues and co-authors (Bradley and Hughes) of Professor Mann admitted in their independent (i.e., without Prof. Mann as co-author) and updated publication, "A caveat to [our] conclusion [about northern hemisphere temperature change over the last 1000 years] is that *the current proxy-based reconstructions do not extend to the end of the 20th century, but are patched on to the instrumental record of the last 2–3 decades* [emphasis added]. This is necessary because many paleo data sets were collected in the 1960's and 1970's, and have not been

[1] Which was awarded the 1989 nation-wide IEEE Nuclear and Plasma Sciences Society Graduate Scholastic Award and the 1991's Rockwell Dennis Hunt Scholastic Award for "the most representative PhD thesis work" at the University of Southern California.

up-dated [NOTE: this statement by Bradley *et al.* (2003) referred primarily to the tree-ring data base from the International Tree-Ring Data base.], so a direct proxy-based comparison of the 1990's with earlier periods is not yet possible." [p. 116 of Bradley *et al.*, 2003, In: Alverson, K., R.S. Bradley and T.F. Pedersen (eds.) *Paleoclimate, Global Change and the Future*. Springer Verlag, Berlin, 105–149]

Agreeing with discussion on p. 260–261 of Soon *et al.* (2003), Bradley *et al.* (2003) cautioned that "in the case of tree rings from some areas in high latitudes, the decadal time-scale climatic relationships prevalent for most of this century appear to have changed in recent decades, possibly because increasing aridity &/or snowcover changes at high latitudes may have already transferred the ecological responses of trees to climate (cf. Jacoby and D'Arrigo 1995; Briffa *et al.* 1998). For example, near the northern tree limit in Siberia, this changing relationship can be accounted for by a century-long trend to greater winter snowfall. This has led to delayed snowmelt and thawing of the active layer in this region of extensive permafrost, resulting in later onset of the growing season (Vaganov *et al.* 1999). It is not yet known how widely this explanation might apply to the other regions where partial decoupling has been observed, but regardless of the cause, it raises the question as to whether there might have been periods in the past when the tree ring-climate response changes, and what impact such changes might have on paleotemperature reconstructions based largely on tree ring data." (p. 116–117).

Bradley *et al.* (2003) also worried that "Paleoclimate research has had a strong northern hemisphere, extra-tropical focus (but even there the record is poorly known in many areas before the 17th century). There are very few high resolution paleoclimatic records from the tropics, or from the extra-tropical southern hemisphere, which leaves many questions (such as the nature of climate in Medieval times) unanswered." (p. 141). Bradley *et al.* continued "All large-scale paleotemperature reconstructions suffer from a lack of data at low latitudes. In fact, most "northern hemisphere" reconstructions do not include data from the southern half of the region (i.e. [missing comma] areas south of 30N). Furthermore, there are so few data sets from southern hemisphere that it is not yet possible to reconstruct a meaningful "global" record of temperature variability beyond the period of instrumental records. For the northern hemisphere records, it must be recognized that the errors estimated for the reconstructions of Mann *et al.* (1999) and Briffa *et al.* (2001) are minimum estimates, based on the statistical uncertainties inherent in the methods used. These can be reduced by the use of additional data (with better spatial representation) that incorporate stronger temperature signals. However, there will always be additional uncertainties that relate to issues such as the constancy of the proxy-climate function over time, and the extent to which modern climate modes (i.e., those that occurred during the calibration interval) represent the full range of climate variability in the past [i.e., similar unresolved research questions had been raised in p. 239–242 and p. 258–264 of Soon *et al.* 2003]. There is evidence that in recent decades some high latitude trees no longer capture low frequency variability as well as in earlier decades of the 20th century (as discussed below in Section 6.8) which leads to concerns over the extent to which this may have also been true in the more distant past. If this was a problem (and currently we are not certain of that) it could result in an inaccurate representation of low frequency temperature changes in the past. Similarly, if former climates were characterized by modes of variability not seen in the calibration period, it is unlikely that the methods now in use would reconstruct those intervals accurately. It may be possible to constrain these uncertainties through a range of regional studies (for example, to examine modes of past variability) and by calibration over different time intervals, but not all uncertainty can be eliminated and so current margins of error must be considered as minimum estimates [meaning the actual range of error is larger than shown in Mann *et al.* 1999 or the IPCC TAR's charts]." (p. 114–115).

It is also very important to heed warnings and cautions from other serious researchers about not over stating the true confidence of a reconstructed climatic result based on indirect proxies. Esper *et al.* (2003, Climate Dynamics, vol. 21, 699–706) modestly apprised of the current situation in reconstructing long-term climatic information from tree rings: "Although these long-term trends agree well with ECS [i.e., Esper, Cook, Schweingruber in 2002, Science, vol. 295, 2250–2253], the amplitude of the multi-centennial scale variations is, however, not understood. This is because (1) no single multi-centennial scale chronology could be built that is not systematically biased in the low frequency domain, and (2) no evidence exists that would support an estimation of the biases either in the LTM [Long-term mean standardization] nor in the RCS [Regional curve standardization] multi-centennial chronologies. Consequently, we also avoided providing formal climate calibration and verification statistics of the chronologies. Note also that the climate signal of the chronologies' low frequency component could not be statistically verified anyway.

This is because the high autocorrelations, when comparing lower frequency trends, significantly reduce the degrees of freedom valid for correlation analyses. We believe that a formal calibration/verification/transfer function approach would leave the impression that the long-term climate history for the Tien Shan [i.e., the location of Esper and five colleagues' study] is entirely understood, which is not the case. Further research is needed to estimate the amplitude of temperature variation in the Alai Range [south of Kirghizia] over the last millennium." (p. 705)

Question 4. Do you claim that the Mann study does not reconstruct regional patterns of temperature change in past centuries?

Response. In Soon *et al.* (2003, Energy & Environment, vol. 14, 233–296), I and my colleagues cautioned that the regional temperature patterns resulted from Mann and colleagues' methodology are too severely restricted by the calibration particular, we are concerned that the regional (and hence larger spatial-scale averages) variability of temperature on multidecadal and centennial time scales deduced from such a method will be underestimated.

Recently, the methodology of Mann *et al.* (1998) has been seriously challenged by McIntyre and McKitrick (2003, Energy & Environment, vol. 14, 751–771) in that "poor data handling, obsolete data and incorrect calculation of principal components" were shown as the errors and defects of Mann *et al's.* paper. The exchange between Mann and colleagues and McIntyre and McKitrick is ongoing, but the use of obsolete data is a clear case of misrepresentation of regional basis of change in Mann et al's work. Further problems in Mann *et al.* (1998) are outlined under Question No. 13 below. Additional documentation (including responses by Prof. Mann and his colleagues) and updates can be found in *http://www.uoguelph.ca/rmckitri/research/trc.html.*

Question 5. Do you maintain that the Mann study extrapolated global temperature estimates from the northern hemisphere?

Response. I have not seen any global temperature curves presented in the two earlier studies by Mann *et al.* (1998 and 1999). But please consider the deep concerns about the lack of proxy data especially over the tropics (30N to 30S) and the southern hemisphere raised by Soon *et al.* (2003) and even in the independent paper by Professor Mann's close colleagues and co-authors (Bradley and Hughes), i.e., in Bradley *et al.* (2003), discussed under Question No. 3 above.

"Global" temperature estimates, based on indirect climate proxies, from 200–1980 were shown in Mann and Jones (2003, *Geophysical Research Letters*, vol. 30 (15), 1820) as Figure 2c. But I am unsure if the temperature series presented by Mann and Jones (2003) could adequately represents the variability over the whole globe since it was openly admitted that the proxies used covered only 8 "distinct regions" in the Northern Hemisphere and 5 for the Southern Hemisphere (see the coverage of proxies shown in Figure 1 of Mann and Jones, 2003).

More importantly, Soon *et al.* (2004, *Geophysical Research Letters*, vol. 31, L03209) showed that the 40-year smoothed instrumental temperature trend for the Northern Hemisphere shown as Figure 2a of Mann and Jones (2003) has a physically implausible high value at year 2000 (see more discussion in Question No. 6 below). We caution that the extremely rapid rate of warming trend of 1 to 2.5° C per decade implied by the published results by Mann and his colleagues over the last one to 2 years [comparing Mann and Jones (2003) with both Mann (2002, Science, vol. 297, 1481–1482) and Mann *et al.* (2003, *Eos*, 84(27), 256–257)], is most likely due to the artifacts of methodology and their procedure of trend smoothing. I am submitting the pdf file (SLB-GRL04-NHtempTrend.pdf) of Soon *et al.* (2004) for the record of the committee.

Question 6. Do you maintain that historical and instrumental temperature records that are available indicate colder northern hemisphere temperature conditions than the Mann et al northern hemisphere temperature reconstruction in the past centuries?

Response. I am not sure about the meaning of this question. But when contrasted with borehole-based reconstruction, the Northern Hemisphere terrestrial temperatures produced by Mann *et al.* (1998, 1999) over the last 500 years may have been too warm by about 0.4)° C during the 17th–18th century (see Huang *et al.* 2000, Nature, vol. 403, 756–758). Recent attempts by Mann *et al.* (2003, *Journal of Geophysical Research*, vol. 108. (D7), 4203) and Mann and Schmidt (2003, *Geophysical Research Letters*, vol. 30 (12), 1607) to rejustify and defend the Mann *et al.* (1998, 1999) results have been shown to be either flawed or invalid by Chapman *et al.* (2004, *Geophysical Research Letters*, vol. 31, L07205) and by Pollack and Smerdon (2003, Geophysical Research Abstract of EGS, vol. 6, 06345). The eventual fact will no doubt emerge with increased understanding, but Chapman *et al.* (2004) warned

that "A second misleading analysis made by Mann and Schmidt [2003] concerns use of end-points in reaching a numerical conclusion. . . . It is based on using end points in computing changes in an oscillating time series, and is just bad science."

With regard to instrumental thermometer data of the past 100–150 years, it is important to note that Soon *et al.* (2004) has recently shown that the 40-year smoothed Northern Hemisphere temperature trend shown in Mann and Jones (2003) has a physically implausible high value at the year 2000 endpoint especially when studied in context with previous published results by Mann *et al.* (2003, *Eos*, vol. 84 (27), 256–257) and Mann (2002, Science, vol. 297, 1481–1482). This important updated information, admittedly with the benefit of hindsight, together with the works by Chapman *et al.* (2004) and McIntyre and McKitrick (2003), showed clearly that the Northern Hemisphere temperature trends, either proxy-based or instrumental, derived by Mann *et al.* (1998, 1999) and Mann and Jones (2003) are not reliable.

Question 7. Is it your understanding that during the mid-Holocene optimum period (the period from 4000–7000 B.C.) that annual mean global temperatures were more than a degree C warmer than the present day?

Response. Again, I am not sure if there are sufficient proxy data that would allow a meaningful quantitative estimate of annual mean global temperatures back six to nine thousand years. But in a new paper for the Quaternary Science Reviews, Darrell Kaufman and 29 co-authors (2004, Quaternary Science Reviews, vol. 23, 529–560) found that indeed there are clear evidence for warmer than present conditions during the Holocene at 120 out of 140 sites they compiled across the Western Hemisphere of the Arctic. Kaufman *et al.* (2004) estimated that, at the 16 terrestrial sites where quantitative data are available, the local Holocene Thermal Maximum summer temperatures were about 1.6±0.8° C higher than the average of the 20th century. The coarse temperature map sketched on the NOAA's Paleoclimatology web site: *http://www.ngdc.noaa.gov/paleo/globalwarming/images/polarbigb.gif* suggests that the summer temperatures 6000 years ago may have been 2 to 4° C warmer than present in the other sector (Eastern Hemisphere) of the Arctic.

Question 8. As a climatologist, can you explain what kind of quantitative analysis it takes to determine whether or not the last 50 years has been unusually warm compared to the last 1000 years?

Response. The theoretical requirement is fairly simple: (a) find local and regional proxies that are sensitive to variations of temperature on timescales of decade, several decades and century; (b) have sufficient spatial coverage of these local and regional proxies. Then one would be able to compare the last 50 years of the 1000-year record with the previous 950 years.

Soon *et al.* (2003) had indeed initiated an independent effort in this direction and concluded that a truly global or hemispheric averaged temperature record for the past 1000 years is not yet forthcoming because of the large and disparate range of the indirect local and regional proxies to temperature such that a robust ability of different proxies in capturing all the necessary scales of variability cannot yet be confirmed. The main problem I foresee in having any definitive answers for now is related to the fact that the statistical association of each proxy to climatic variables like temperature can itself be variable and changing depending on the location and time interval. But I am not sure if the sole focus on temperature as the measure of "climate" is sensible if not unnecessarily narrow.

In Soon *et al.* (2003), we consider climate to be more than just temperature so we did not narrowly restrict ourselves to only temperature-sensitive proxies. For example, in addition to temperature, we are equally concerned about expansion and reduction of forested and desert-prone areas, tree-line growth limit, sea ice changes, balances of ice accumulation and ablation in mountain glaciers and so on. When studying the ice balance for a glacier, it is important to insist that although glaciers are very important indicators of climate change over a rather long time-scale, they are not simply thermometers as often confused by heated discussion pointing to evidence for global warming by carbon dioxide (see additional discussion on factors, especially atmospheric carbon dioxide, in determining Earth's climate and its change under Questions No. 19, 20, 25, 30 and 35 below). Examples include statements by Will Stefen, director of the International Geosphere-Biosphere Program, "Tropical glaciers are a bellweather of human influence on the Earth system" (quoted in the article "The melting snows of Kilimanjaro" by Irion, 2001, Science, vol. 291, 1690–1691) or by Professor Lonnie Thompson, Ohio State University,

> "We have long predicted that the first signs of changes caused by global warming would appear at the few fragile, high-altitude ice caps and glaciers within the tropics . . . [t]hese findings confirm those predictions. We need to take the first steps to reduce carbon dioxide emissions. We are currently doing

> nothing. In fact, as a result of energy crisis in California—and probably in the rest of the country by this summer—we will be investing even more in fuel-burning power plants. That will put more power in the grid but, at the same times it will add carbon dioxide to the atmosphere, amplifying the problem" (quoted in Ohio State University's press release, *http://www.acs.ohio-state.edu/units/research/archive/glacgone.htm*).

A clarification about the physical understanding of modern glacier retreats and climate change, especially those on Kilimanjaro, is necessary and has been forthcoming with important research progress. First, Molg *et al.* (2003, *Journal of Geophysical Research*, vol. 108 (D23), 4731) recently concluded that their study:

> "highlights that modern glacier retreat on Kilimanjaro is much more complex than simply attributable to 'global warming only', a finding that conforms with the general character of glacier retreat in the global tropic [Kaser, 1999]: a process driven by a complex combination of changes in several different climatic parameters . . . with humidity-related variables dominating this combination."

In another new paper for the International Journal of Climatology, Kaser *et al.* (2004, International Journal of Climatology, "Modern glacier retreat on Kilimanjaro as evidence of climate change: Observations and facts", vol. 24, 329–339; available from *http://geowww.uibk.ac.at/glacio/LITERATUR/index.html*) provided clear answers that neither added longwave radiation from a direct addition of atmospheric CO_2 nor atmospheric temperature were the key variables for the observed changes, as revealed in this long but highly informative passage:

> "Since the scientific exploration of Kilimanjaro began in 1887, when Hans Meyer first ascended the mountain (not to the top at this time, but to the crater rim), a central theme of published research has been the drastic recession of Kilimanjaro's glaciers (e.g., Meyer, 1891, 1900; Klute, 1920; Gilman, 1923; Jager, 1931; Geilinger, 1936; Hunt, 1947; Spink, 1949; Humphries, 1959; Downie and Wilkinson, 1972; Hastenrath, 1984; Osmastion, 1989; Hastenrath and Greischar, 1997). Early reports describe the formation of notches, splitting up and disconnection of ice bodies, and measurements of glacier snout retreat on single glaciers, while later books and papers advance to reconstructing glacier surface areas. . . . Today, as in the past, Kilimanjaro's glaciers are markedly characterized by features such as penitentes, cliffs (Figure 3a/b) [not reproduced here], and sharp edges, all resulting from strong differential ablation. These features illustrate the absolute predominance [emphasis added] of incoming shortwave radiation and latent heat flux in providing the energy for ablation (Kraus, 1972). A positive heat flux from either longwave radiation or sensible heat flux, if available, would round-off and destroy the observed features within a very short time ranging from hours to days. On the other hand, if destroyed, the features could only be sculptured again under very particular circumstances and over a long time. Thus, the existence of these features indicates that the present summit glaciers are not experiencing ablation due to sensible heat (i.e., from positive air temperature). Additional support for this is provided by the Northern Icefield air temperature recorded from February 2000 to July 2002, which never exceeded $-1.6°$ C, and by the presence of permafrost at 4,700 m below Arrow Glacier on the western slope . . ."

Kaser *et al.* (2004) continue with this "synopsis of interpretations and facts":

> "A synopsis of (i) proxy data indicating changes in East African climate since ca. 1850, (ii) 20th century instrumental data (temperature and precipitation), and (iii) the observations and interpretations made during two periods of fieldwork (June 2001 and July 2002) strongly support the following scenario. Retreat from a maximum extent of Kilimanjaro's glaciers started shortly before Hans Meyer and Ludwig Purtscheller visited the summit for the first time in 1889, caused by an abrupt climate change to markedly drier conditions around 1880. Intensified dry seasons accelerated ablation on the respectively illuminated vertical walls left in the hole on top by Reusch Crater as a result of *volcanic activity* [emphasis added]. The development of vertical features may also have started on the outer margins of the plateau glaciers before 1900, primarily as the formation of notches, as explicitly reported following field research in 1898 and 1912 (Meyer, 1900; Klute, 1920). A current example of such a notch development is the hole in the Northern Icefield (see Figure 2). Once started, the lateral retreat was unstoppable, maintained by solar radiation despite less negative mass balance conditions on horizontal glacier surfaces, and will come to an end only when the glaciers on the summit plateau have disappeared. This is most probable within next decades, if the trend revealed in Figure 1 continues. Positive air temperatures have not contributed to the recession process

on the summit so far. The rather independent slope glaciers have retreated far above the elevation of their thermal readiness, responding to dry conditions. If present precipitation regime persists, these glaciers will most probably survive in positions and extents not much different from today. This is supported by the area determinations in Thompson's *et al.* (2002) map, which indicate that slope glaciers retreated more from 1912 to 1952 than since then. From a hydrological point of view, melt water from Kibo's glaciers has been of little importance to the lowland in modern times. Most glacier ablation is due to sublimation, and where ice does melt it immediately evaporates into the atmosphere. Absolutely no signs of runoff can be found on the summit plateau, and only very small rivers discharge from the slope glaciers. Rainfall reaches a maximum amount at about 2,500 m a.s.l. [above sea level] (Coutts, 1969), which primarily feeds the springs at low elevation on the mountain; one estimate attributes 95 percent of such water to a forest origin (Lambrechts *et al.*, 2002). The scenario presented offers a concept that implies climatological processes *other than increased air temperature* [emphasis added] govern glacier retreat on Kilimanjaro in a direct manner. However, it does not rule out that these processes may be linked to temperature variations in other tropical regions, e.g., in the Indian Ocean (Latif *et al.*, 1999; Black *et al.*, 2003)."

Lindzen (2002, Geophysical Research Letter, vol. 29, paper 2001GL014360) further added that "Recent papers show that deep ocean temperatures have increased somewhat since 1950, and that the increase is compatible with predictions from coupled GCMs [General Circulation Models]. The inference presented is that this degree of compatibility constitutes a significant test of the models. . . . [But] it would appear from the present simple model (which is similar to what the IPCC uses to evaluate scenarios) that the ocean temperature change largely reflects only the fact that surface temperature change is made to correspond to observations, and says almost nothing about model climate sensitivity. . . . It must be added that we are dealing with observed surface warming that has been going on for over a century. The oceanic temperature change [at depth of 475 m or so] over the period reflects earlier temperature change at the surface. How early depends on the rate at which surface signals penetrate the ocean." In other words, the recently noted warming of the deeper ocean is not a proof of global surface and atmospheric warming by increasing CO_2 in the air because the parameters of climate sensitivity and rate of ocean heat uptake are not sufficiently well quantified. In addition, if the earlier oceanic surface temperature warming mentioned by Lindzen were indeed initiated and occurred substantially long ago, then there would be no association of that change to man-made CO_2 forcing.

Question 9. The IPCC has found that the late 20th century is the warmest period in the last 1000 years, for average temperature in the northern hemisphere. Does your paper provide a quantitative analysis of average temperatures for the northern hemisphere for this specific time period—that is, for the later half of the 20th century?

Response. It should be understood that (1) the conclusion of the IPCC Working Group I's Third Assessment Report (2001; TAR), (2) the evidence shown in Figure 1b of the Summary for Policymaker, (3) Figure 5 of the Technical Summary, and (4) Figure 2.20 in Chapter 2 of TAR were all derived directly from the conclusion in Mann *et al.* (1999) and Figure 3a of Mann *et al.* (1999). Therefore all comments and criticisms presented in this Q&A about Mann *et al.* (1999) apply to the IPCC TAR's conclusion. In addition, Soon *et al.* (2004) recently cautioned that the 40-year smoothed northern hemisphere temperature trend shown in Figure 2.21 of TAR (2001) cannot be replicated according to the methodology described in the caption of Figure 2.21. The failure in replication introduces a significant worry about the actual quality of scientific efforts behind the production of Figure 2.21 in TAR (2001).

The answer to the second part of your direct question is no. Here are the related reasons why a confident estimate of the averaged northern hemisphere temperature for the full 1000 years (including the full 20th century) is not yet possible, despite what had been claimed by Mann *et al.* (1999). First, several authors, including those detailed in section 5.1 of Soon *et al.* (2003) and those pointed out in Question No. 6, had shown that the 1000-year series of mathematical temperature derived by Mann *et al.* (1999) has significantly underestimated the multidecadal and centennial scale changes. Second, the focus of Soon *et al.* (2003) is to derive understanding of climatic change on local and regional spatial scales, instead of over the whole northern hemisphere per se, because those are the most relevant measures, in practical sense, of change. In addition, we provided the first-order attempt to collect all available climate proxies relevant for local and regional climatic changes, but not re-

stricted to temperature alone. But more pertinent to your question is the fact discussed in Soon *et al.* (2003) that different proxies respond with differing sensitivities to different climatic variables, seasons, plus spatial and temporal scales, so that a convenient derivation of a self-consistent northern hemisphere averaged annual mean temperature for the full 1000 years, desirable as the result may be, is not yet possible.

Question 10. Does your paper provide any quantitative analysis of temperature records specifically for the last 50 years of the 20th century?

Response. Soon *et al.* (2003) considered all available proxy records with no particular prejudice. If the individual proxy record covers up to the last 50 years of the 20th century, then quantitative comparisons are performed, mostly according to the statements from the original authors. Please consider some of the detailed quantitative discussion in section 4 of Soon *et al.* (2003) and the qualitative results compiled in Table 1 of that paper.

Question 11. In an article in the Atlanta Journal Constitution (June 1, 2003), you were quoted as acknowledging during a question period at a previous Senate luncheon that your research does not provide a comprehensive picture of the Earth's temperature record and that you questioned whether that is even possible, and that you did not, ". . . see how Mann and the others could 'calibrate' the various proxy records for comparison." How then does your analysis provide a comprehensive picture of Earth's temperature record or have any bearing on the finding by the IPCC, that the late 20th century is the warmest in the last 1000 years?

Response. Thank you for referencing the article. I must first state on the record that contrary to the claim in this Atlanta Journal Constitution (June 1, 2003) article *http://www.ajc.com/business/content/business/0603/01warming.html,* the writer, never, as claimed, conducted a telephone interview with me. No such conversation took place and I am rather shocked by this false claim. This fact has gone uncorrected until now.

The strengths and weaknesses of my research works are fully discussed in Soon *et al.* (2003). The paper documented detailed local and regional changes in several climatic variables to try to obtain a broader understanding of climate variability. We concluded that:

> "Because the nature of the various proxy climate indicators are so different, the results cannot be combined into a simple hemispheric or global quantitative composite. However, considered as an ensemble of individual observations, an assemblage of the local representations of climate establishes the reality of both the Little Ice Age and the Medieval Warm Period as climatic anomalies with worldwide imprints, extending earlier results by Bryson *et al.* (1963), Lamb (1965), and numerous other research efforts. Furthermore, these individual proxies are used to determine whether the 20th century is the warmest century of the 2nd Millennium at a variety of globally dispersed locations. Many records reveal that the 20th century is likely not the warmest nor a uniquely extreme climatic period of the last millennium, although it is clear that human activity has significantly impacted some local environments."

The question on the difficult problem of calibrating proxies of differing types and sensitivities to climatic variables is discussed in Soon *et al.* (2003) and some criticisms on the weaknesses of the reconstruction by Mann *et al.* (1999) or the related IPCC TAR's conclusion are listed especially under Questions No. 6 and 9.

Question 12. Do you believe that appropriate statistical methods do not exist for calibrating statistical predictors, including climate proxy records, against a target variable, such as the modern instrumental temperature record?

Response. True progress in the field of paleoclimatology will certainly involve a better and more robust means of interpreting and quantifying the variations and changes seen in each high-resolution proxy record. The issue is not merely a problem awaiting solution through appropriate statistical methods like the EOF methodology adopted by Mann *et al.* (1998, 1999). On pp. 241–242 of Soon *et al.* (2003), we briefly outlined our straight-forward approach and contrasted it to the one used by Mann and colleagues that does not necessarily lead to results with physical meaning and reality.

Question 13. In determining whether the temperature of the "Medieval Warm Period" was warmer than the 20th century, does your study analyze whether a 50-year period is either warmer or wetter or drier than the 20th century? If so, why is it appropriate to use indicators of drought and precipitation directly to draw inferences of past temperatures? Please list peer-reviewed works that specifically support the use of these indicators for inferring past temperature.

Response. The detailed discussion behind our usage of the term "Medieval Warm Period" or "Little Ice Age" was described in Soon *et al.* (2003). We are mindful that the two terms should definitely include physical criteria and evidence from the thermal field. But we emphasize that great bias would result if those thermal anomalies were dissociated from hydrological, cryospheric, chemical, and biological factors of change. So indeed our description of a Medieval Climatic Anomaly (see a similar sentiment later reported by Bradley *et al.* 2003, Science, vol. 302, 404–405) in Soon *et al.* (2003) includes a warmer time that contains both drought or flooding conditions depending on the locations.

With regard to the last part of your question, I would answer by detailing only one example—Mann et al (1998). This influential study used both direct precipitation measurements and precipitation proxies as temperature indicators. This study was indeed applied by the IPCC TAR (2001). These include historical precipitation measurements in 11 grid cells, two coral proxies (reported in Mann *et al.* [1998] as precipitation proxies; see *http://www.ngdc.noaa.gov/paleo/ei/data—supp.html* for this and following references), two ice core proxies, 3 reconstructions of spring precipitation in southeast United States by Stahle and Cleaveland from tree ring data, 12 principal component series for tree rings in southwestern United States and Mexico reported as precipitation proxies by Stahle and Cleaveland (and Mann *et al.* 1998) and one tree ring series in Java—making a total of 31 precipitation series used as proxies in temperature reconstruction by Mann *et al.* (1998). In this peer-reviewed article, for the precipitation data in a grid cell in New England, the researchers apparently used historical data from Paris, France (please see Figure 2 of McIntyre and McKitrick, 2003 and their discussion on pp. 758–759). For a grid cell near Washington DC, the researchers used historical data from Toulouse, France. For a grid cell in Spain, the researchers used precipitation data from Marseilles, France. Of the 11 precipitation series used in Mann *et al.* (1998), only one series (Madras, India) is correctly located. The precipitation data used by these authors cannot be identified in the source cited in paper Mann *et al.* (1998). While precipitation data and precipitation-related proxies can be instructive in providing information on past distribution of moisture and circulation patterns (and thus temperature), it is important to correctly identify the series used and important not to use data from the wrong continent for historical reconstructions.

Question 14. Do you maintain that any two 50-year periods that occur within a multi-century interval can be considered 'coincident' from a climatic point of view?

Response. The question raised here about the connection of any two 50-year periods in any two regions to be related from climatic point of view is both important and interesting. But the answer will be strongly dependent on the nature of forcings and feedbacks involved. If longer-term cryospheric or oceanic processes are involved then the answer would be yes.

Question 15. Do your two recent studies employ an analysis (that is, a statistical or analytical operation performed upon numerical data) of a single proxy climate record?

Response. The meaning of this question is not entirely clear to me. But I would say yes under the context of what is being said.

Question 16. Has your study produced a quantitative reconstruction of past temperature patterns? Do you have a measure of uncertainty or verification in your description of past temperatures?

Response. The results and conclusion of Soon *et al.* (2003) are best judged by the paper itself. Quantitative assessments of local and regional changes through the climatic proxies are discussed in section 4 of that paper as well as some qualitative picture described in Figures 1, 2 and 3 of that paper. Again, Soon *et al.* (2003) did not tried to distill all the collected proxies down to produce a strict temperature-only result since we are interested in a broader understanding of climate variability. Part of the answers given under Questions No. 9 and 11 can help elaborate what was done by Soon *et al.* (2003). I would also like to direct your attention to the two warnings listed under Question No. 3 by Bradley *et al.* (2003) and Esper *et al.* (2003) concerning any undue, over confidence in promoting quantitative certainties in the reconstruction of past temperatures through highly imprecise black boxes of indirect proxies.

Question 17. Your study indicates that you have compiled the results of hundreds of previous paleo-climate studies. Have you verified your interpretation of the hundreds of studies with any of the authors/scientists involved in those studies? If so, how many?

Response. Specific authors and scientists that provided help in our work were listed in the acknowledgement section (p. 272) of Soon *et al.* (2003). We have also re-

ceived generous help and comments from several scientists who are certainly highly qualified in terms of paleoclimatic studies. But the ultimate quality and soundness of our research shall always be our own responsibility.

In the September 5, 2003 Chronicle for Higher Education article (by Richard Monastersky), there were indeed two very serious accusations that suggested that Soon *et al.* (2003) had misrepresented or abused the conclusions by two original authors whose work we had cited. Our corrections and explanations to these unfortunately false claims can be studied from the documentation listed in the URL *http://cfa-www.harvard.edu/*~wsoon/ChronicleHigherEducation03-d (read especially Sep12-lettoCHE3.doc and Sep12-lettoCHE4.doc).

Question 18. What was earth's climate like the last time that atmospheric concentrations of carbon dioxide were at today's levels or about 370 parts per million (ppm) and what were conditions like when concentration were at 500 ppm, which will occur around 2060 or so?

Response. Co-answer to this question is listed under Question No. 19 below.

Question 19. Please describe any known geologic precedent for large increases of atmospheric CO_2 without simultaneous changes in other components of the carbon cycle and the climate system.

Response. My July 29, 2003 testimony was about the climate history of the past 1000 years detailed in Soon *et al.* (2003) rather than any potential (causal or otherwise) relationship between atmospheric carbon dioxide and climate. The fact remains that the inner working of the global carbon cycle and the course of future energy use are not sufficiently understood or known to warrant any confident prediction of atmospheric CO_2 concentration at year 2060. Please consider co-answer to this question under Question No. 25 below.

However, it is abundantly obvious that atmospheric CO_2 is not necessarily an important driver of climate change. It is indeed a puzzle that despite the relative low level of atmospheric CO_2 of no more than 300 ppm in the past 320–420 thousand years (Kawamura *et al.*, 2003, Tellus, vol. 55B, 126–137) compared to the high levels of 330–370 ppm since the 1960's there is the clear suggestion of significantly warmer temperatures at both Vostok and Dome Fuji, East Antarctica, during the interglacials at stage 9.3 (about 330 thousand years before present; warmer by about 6° C) and stage 5.5 (about 135 thousand years before present; warmer by about 4.5° C) than the most recent 1000 years (see Watanabe *et al.*, 2003, Nature, vol. 422, 509–512; further detailed discussion on environmental changes in Antartica over the past 1000 years or so, including the most recent 50 years can be found in section 4.3.4 or pp. 256–257 of Soon *et al.* 2003).

But there are important concerns about the retrieval of information on atmospheric CO_2 levels from ice cores. Jaworowski and colleagues (1992, The Science of the Total Environment, vol. 114, 227–284) explained that:

> "Ice is not a rigid material suitable for preserving the original chemical and isotopic composition of atmospheric gas inclusion. Carbon dioxide in ice is trapped mechanically and by dissolution in liquid water. A host of physicochemical processes redistribute CO_2 and other air gases between gaseous, liquid and solid phases, in the ice sheets in situ, and during drilling, transport and storage of the ice cores. This leads to changes in the isotopic and molecular composition of trapped air. The presence of liquid water in ice at low temperatures ['even below—70° C'] is probably the most important factor in the physico-chemical changes. The permeable ice sheet with its capillary liquid network acts as a giant sieve which redistributes elements, isotopes and micro-particles. Carbon dioxide in glaciers is contained: (1) in interstitial air in firn; (2) in air bubbles in ice; (3) in clathrates; (4) as a solid solution in ice crystals; (5) dissolved in intercrystalline veins and films of liquid brine; and (6) in dissolved and particulate carbonates. Most of the CO_2 is contained in ice crystals and liquids, and less in air bubbles. In the ice cores it is also present in the secondary gas cavities, cracks, and in the traces of drilling fluids.
>
> The concentration of CO_2 in air recovered from the whole ice is usually much higher than that in atmospheric air. This is due to the higher solubility of this gas in cold water, which is 73.5- and 35-times higher than that of nitrogen and oxygen, respectively. The composition of other atmospheric gases (N_2, O_2, Ar) is also different in ice and in air inclusions than in the atmosphere. Argon–39 and 85Kr data indicate that 36–100 percent of air recovered from deep Antarctic ice cores is contaminated by recent atmospheric air during field and laboratory processing. Until about 1985, CO_2 concentrations in gas recovered from primary air bubbles and from secondary gas cavities in pre-industrial and ancient ice were often reported to be much higher than in the present atmosphere. After 1985, only concentrations below the current atmospheric level were pub-

lished. Our conclusion is that both these high and low CO_2 values do not represent real atmospheric content of CO_2.

Recently reported concentrations of CO_2 in primary and secondary gas inclusions from deep cores, covering about the last 160,000 years, are much below the current atmospheric level, although several times during this period the surface temperature was 2–4.5° C higher than now. If these low concentrations of CO_2 represented real atmospheric levels, this would mean (1) that CO_2 had not influenced past climatic changes, and (2) that climatic changes did not influence atmospheric CO_2 levels." (p. 272–273)

Additional historical evidence reveals natural occurrences of large, abrupt climatic changes that are not uncommon and they occurred without any known causal ties to large radiative forcing change. Phase differences between atmospheric CO_2 and proxy temperature in historical records are often not fully resolved; but atmospheric CO_2 has shown the tendency to follow rather than lead temperature and biosphere changes (see e.g., Dettinger and Ghil, 1998, Tellus, vol. 50B, 1–24; Fischer *et al.*, 1999, Science, vol. 283, 1712–1714; Indermuhle *et al.*, 1999, Nature, vol. 398, 121–126).

In addition, there have been geological times of global cooling with rising CO_2 (during the middle Miocene about 12.5–14 million years before present [Myr BP], for example, with a rapid expansion of the East Antarctic Ice Sheet and with a reduction in chemical weathering rates), while there have been times of global warming with low levels of atmospheric CO_2 (such as during the Miocene Climate Optimum about 14.5–17 Myr BP as noted by Panagi *et al.*, 1999, Paleocenoragphy, vol. 14, 273–292). A new study of atmospheric carbon dioxide over the last 500 million years (Rothman, 2002, Proceedings of the (US) National Academy of Sciences, vol. 99, 4167–4171) concluded that, "CO_2 levels have mostly decreased for the last 175 Myr. Prior to that point [CO_2 levels] appear to have fluctuated from about two to four times modern levels with a dominant period of about 100 Myr. . . . The resulting signal exhibits no systematic correspondence with geologic record of climatic variations at tectonic time scales."

Question 20. According to a study published in Science magazine, [B. D. Santer, M. F. Wehner, T. M. L. Wigley, R. Sausen, G. A. Meehl, K. E. Taylor, C. Amman, W. M. Washington, J. S. Boyle, and W. Bruggemann Science 2003 July 25; 301: 479–483], manmade emissions are partly to blame for pushing outward the boundary between the lower atmosphere and the upper atmosphere. How does that fit with the long-term climate history and what are the implications?

Response. It should first be noted that Pielke and Chase (2004, Science, vol. 303, 1771b; and see p. 1771c by Santer *et al.* and additional counter-reply by Pielke and Chase, with input from John Christy and Anthony Reale, available as paper 278b at *http://blue.atmos.colostate.edu/publications/reviewedpublications.shtml*) had criticized and challenged Santer *et al.*'s claim and conclusion that,

"[o]ur results are relevant to the issue of whether the 'real-world' troposphere has warmed during the satellite era. . . . The direct evidence is that in the ALL experiment [i.e., climate model results that included changes in well-mixed greenhouse gases, direct scattering effects of sulfate aerosols, tropospheric and stratospheric ozone, solar total irradiance and volcanic aerosols; see more discussion below], the troposphere warms by 0.07° C/decade over 1979–1999. This warming is predominantly due to increases in well-mixed greenhouse gases. . . . Over 1979–1999, roughly 30 percent of the increase in tropopause height in ALL is explained by greenhouse gas-induced warming of the troposphere. Anthropogenically driven tropospheric warming is therefore an important factor in explaining modeled changes in tropopause height."

In contrast, Pielke and Chase (2004) offered the observed evidence and concluded that

"[g]lobally averaged tropospheric temperature trends are statistically indistinguishable from zero. Thus, the elevation of the globally averaged tropopause report in [Santer *et al.*, 2003] cannot be attributed to any detectable tropospheric warming over this period." In addition, "the climate system is much more complex than defined by tropospheric temperature and tropopause changes. Linear trend analysis [in Santer *et al.*, 2003] is of limited significance. Changes in global heat storage provide a more appropriate metric to monitor global warming than temperature alone."

Soon and Baliunas (2003, Progress in Physical Geography, vol. 27, 448–455) had also previously outlined the incorrect fingerprint of CO_2 forcing observed in even the best and sophisticated version of climate models thus far. A more general and comprehensive discussion about the fundamental difficulties on modeling the effects of

carbon dioxide using current generation of climate models is given in Soon *et al.* (2001, Climate Research, vol. 18, 259–271). Thus, the new paper by Santer *et al.* (2003) does not supercede or overcome the difficulties with respect to General Circulation Climate Models raised in Soon and Baliunas (2003).

Both the meaning and strength of the model-dependent results shown in Santer *et al.* (2003) remain doubtful and weak for several additional reasons.

First, Figure 2 of Santer *et al.* (2003) itself confirmed that the modeled changes in tropopause height are caused mainly by large stratospheric cooling related to changes in stratospheric ozone (they admitted so even though their note No. 35 indicates that their numerical experiments did not separate tropospheric and stratospheric ozone changes) rather than by the well-mixed greenhouse gases that are supposed to be the subject of concern. Second, the model experiments of Santer *et al.* (2003) did not include changes in stratospheric water vapor which is known to be a significant factor for the observed stratospheric cooling (see e.g., Forster and Shine, 1999, *Geophysical Research Letters*, vol. 26, 3309–3312). Third, the failure to account for stratospheric water vapor contradicted the documented significant increases of stratospheric water vapor in the past half-century from a variety of instrumentations (e.g., Smith et al, 2000, *Geophysical Research Letters*, vol. 27, 1687–1690; Rosenlof *et al.*, 2001, *Geophysical Research Letters*, vol. 28, 1195–1198; though Randel *et al.* [2004, Journal of the Atmospheric Sciences, submitted] recently noted that unusually low water vapor has been observed in the lower stratosphere for 2001–2003). Fourth, the model experiments by Santer *et al.* (2003) had clearly neglected (see note No. 18 of that paper) the role of the Sun's ultraviolet radiation that is not only known to be variable (e.g., Fontenla *et al.* 1999, The Astrophysical Journal, vol. 518, 480–499; White *et al.*, 2000, Space Science Reviews, vol. 94, 67–74) but also known to exert important influence on both the chemistry and thermal properties in the stratosphere and troposphere (e.g., Larkin *et al.*, 2000, Space Science Reviews, vol. 94, 199–214).

Finally, the physical representation of aerosol forcing (which should not be restricted to sulfate alone) in Santer *et al.* (2003) is clearly not comprehensive and at best highly selective. Early on, Russell *et al.* (2000, *Journal of Geophysical Research*, vol. 105, 14891–14898) cautioned that

> "[o]ne danger of adding aerosols of unknown strength and location is that they can be tuned to give more accurate comparisons with current observations but cover up model deficiencies."

Anderson *et al.* (2003, Science, vol. 300, 1103–1104 and see also exchanges in Crutzen *et al.*, 2003, vol. 303, 1679–1681) recently cautioned that:

> "we argue that the magnitude and uncertainty of aerosol forcing may affect the magnitude and uncertainty of total forcing [i.e., 'the global mean sum of all industrial-era forcings'] to a degree that has not been adequately considered in climate studies to date. Inferences about the causes of surface warming over the industrial period and about climate sensitivity may therefore be in error. . . . Unfortunately, virtually all climate model studies that have included anthropogenic aerosol forcing as a driver of climate change (diagnosis, attribution, and projection studies; denoted 'applications' in the figure) have used only aerosol forcing values that are consistent with the inverse approach. If such studies were conducted with the larger range of aerosol forcings determined from the forward calculations, the results would differ greatly. The forward calculations raise the possibility that total forcing from preindustrial times to the present . . . has been small or even negative. If this is correct, it would imply that climate sensitivity and/or natural variability (that is, variability not forced by anthropogenic emissions) is much larger than climate models currently indicate. . . . In addressing the critical question of how the climate system will respond to this [anthropogenic greenhouse gases'] positive forcing, researchers must seek to resolve the present disparity between forward and inverse calculations. *Until this is achieved, the possibility that most of the warming to date is due to natural variability, as well as the possibility of high climate sensitivity, must be kept open.* [emphasis added]"

To further understand the complexity of calculating aerosol forcing, Jacobson (2001, *Journal of Geophysical Research*, vol. 106, 1551–1568) has to account for a total of 47 species "containing natural and/or anthropogenic sulfate, nitrate, chloride, carbonate, ammonium, sodium, calcium, magnesium, potassium, black carbon, organic matter, silica, ferrous oxide, and aluminium oxide" in his recent estimate of only the global direct radiative forcing by aerosols. (Jacobson [2001] found that the global direct radiative forcing by anthropogenic aerosols is only -0.12 W/m^2 while the forcing by combined natural and anthropogenic sources is -1.4 W/m^2.) There are also the indirect aerosol effects. Temperature or temperature change is

clearly not the only practical measure of effects by aerosols. Haywood and Boucher (2000, Reviews of Geophysics, vol. 38, 513–543) stressed the fact that the indirect radiative forcing effect of the modification of cloud albedo by aerosols could range from -0.3 to -1.8 W/m^2, while the additional aerosol influences on cloud liquid water content (hence, precipitation efficiency), cloud thickness and cloud lifetime are still highly uncertain and difficult to quantify (see e.g., Rotstayn and Liu, 2003, *Journal of Climate*, vol. 16, 3476–3481). This is why one can easily appreciate the difficulties faced by Santer *et al.* (2003) because climate forcing by aerosols is not only known within a wide range of uncertainties but also to a large degree of unknown.

Therefore, I conclude that in addition to the fundamental issues related to climate model representation of physical processes, papers like Santer *et al.* (2003) have also failed the basic requirement for internal consistencies in the accounting for potentially relevant climatic forcing factors and feedbacks. This is why I cannot comment on the implication of this particular study and the meaning of the study for long-term climate history.

Question 21. In your testimony, you discussed there being "warming" and "cooling" for different periods. If you did not construct an integral across the hemisphere or a real timeline, don't your findings really just say there were some warm periods and cool periods, and therefore cannot speak to the issue of the rate of warming or cooling?

Response. I am not sure about the meaning of this question and the quotes. My oral remark was merely referring to "making an accurate forecast that includes all potential human-made warming and cooling effects." The detailed discussion about the climatic and environmental changes for the past 1000 years as deduced from the collection of proxies I had studied was given in Soon *et al.* (2003). I can certainly speak to the rate of warming or cooling at any given location or region when the available proxy, with sufficient temporal resolution, is known or proven to be temperature sensitive.

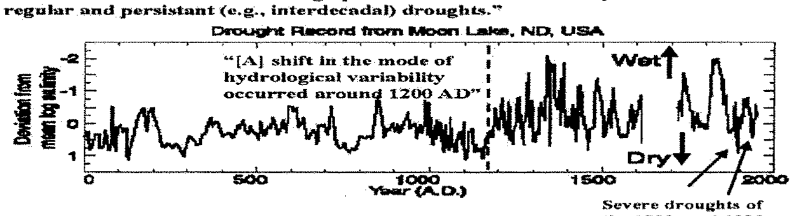
Question 22. Is there any indication that regional climate variations are any larger or smaller at present than over the last 1000 years (with 2003, for example, perhaps being a case with large regional variations from the normal)?

Response. I would not recommend considering the pattern of change from a single year, i.e., 2003, and called it a climate change. But the fact is that in Soon *et al.* (2003) we had carefully studied individual proxy records from various locations and regions. As an example, the 2000-year bottom-sediment record from Moon Lake, North Dakota, shows there is perhaps a distinct shift in the mode of hydrologic variability in the Northern Great Plain region starting around 1200 AD with the more recent period being more variable from the past. But, as indicated in the chart below, the author of this paper also noted that the severe droughts of the 1890's and 1930's around this area are "eclipsed by more extensive droughts before the beginning of the instrumental period."

Greater drought intensity and frequency before AD 1200 in the Northern Great Plain, USA

(Laird et al., 1996, *Nature*, vol. 384, 552-554; Alverson and Oldfield, 2000, PAGES News, vol. 8, no. 1, p. 9)

"The severe droughts of the 1930s and 1890s (positive inferred salinity) are well reconstructed, but are eclipsed by more extensive droughts before the beginning of the instrumental period. There is an abrupt change in drought variability around AD 1200. Before that time, the high plains were characterized by much more regular and persistant (e.g., interdecadal) droughts."



Question 23. In your oral presentation, you talked about "[h]aving computer simulation." Could you please explain what you [as in your original] computer simulation or modeling to which you are referring, and, (a) Has this model gone through the appropriate set of model intercomparison studies like the various othe global models? (b) What forcings have been used to drive it? (c) How does it develop regional climate variations, and are these comparable to observations? and, (d) How does it perform over the 20th century, for example?

Response. I apologize for any potential confusions.

In my oral remark, I said,

> "The entirety of climate proxies over the last 1,000 years shows that over many areas of the world there has been, and continues to be, large climate changes. Those changes provide challenges for the computer simulations of climate. The full models, which explore the Earth region by region, can be tested against the natural patterns of change over the last 1,000 years that are detailed by the climate proxies. Having computer simulations reproduce past patterns of climate, which has been influenced predominantly by natural factors, is key to making an accurate forecast that includes all potential human-made warming and cooling factors."

So in the context of what I said, this question is clearly misdirected by someone who did not understand my remark. I was speaking on the potential application of works like Soon *et al.* (2003) for improving our ability to calculate with confidence the potential effects from man-made factors by first and foremost having a climate model that can at least reproduce some of the observed local and regional changes of the past.

Personally, I am also conducting my research through the help of several climate models (both simple and complex types) appropriate for my interests and I would certainly apply what I found in Soon *et al.* (2003) to my own future studies using climate models. Any additional comments will be beyond the simple context of my oral testimony. But, it may be useful to take note of the comments by Green (2002, Weather, vol. 57, 431–439):

> "It has always worried me that simple models of climate do not seem to work very well. Experts on numerical models say that this is because the atmosphere is very complicated, and that large numerical models and computers are needed to understand it. I worry because I do not know what they have hidden in those models and the programs they use. I wonder what I can compare their models with. Not with each other because they belong to a sort of club, where to have a model that disagrees with everyone else's puts you outside. That is not a bad system, unless of course they are all wrong. Another curiosity of complicated models is that their findings are rarely used to improve the model that preceded them. I would have expected that the more complex model would show where the simpler one had got it wrong, and allow it to be corrected for that misrepresentation."

Question 24. Based on the various comments of your scientific colleagues regarding your paper, including the methodological flaws pointed out in that paper by the former editor-in-chief of Climate Research, are you planning any reworking of your study or any further studies in the paleoclimatic area?

Response. The use of a phrase like "methodological flaws" is a very convenient attempt to dismiss the weight of scientific evidence presented in Soon *et al.* (2003) but unfortunately without any clear nor confirmable basis. Thus far, the only formal criticism of Soon *et al.* (2003) was by Mann *et al.* (2003, *Eos*, vol. 84(27), 256–257) and we had provided our response to that criticism in Soon *et al.* (2003b, *Eos*, vol. 84(44), 473–476). My research interest and work to fully discern and quantitatively describe the local and regional patterns of climate variability over the past 1000 years or so will certainly continue despite this mis-characterization.

It should however not be left unnoticed that several very serious problems in Mann *et al.* (1998, 1999), Mann and Schmidt (2003) and Mann and Jones (2003) had been found recently. Those unresolved anomalies are outlined in my answers to your Questions No. 3, 4, 5, 6, 9 and 13. A careful reworking with a fully open access to all data as well as a fully disclosed transparency of the actual methodologies and detailed applications will be the next important step for paleoclimate reconstruction research.

Question 25. You indicated that there would likely be relatively small climatic response to even substantial increases in the CO_2 concentration. Do you disagree with the radiation calculations that have been done and the trapped energy that they calculate, as per the peer-reviewed literature? If so, please explain.

Response. First, please consider the above discussion on climate forcing factors and climate response sensitivities under Question No. 20 as part of the answers to this question.

Second, I do not believe that I had made any strong claim, one way or another, about the CO_2 forcing and potential response in any specific quantitative term during my testimony (since factually no one can). I do want to comment, as in my response under Question No. 19, that CO_2, as a minor greenhouse gas, is not a determinant of Earth's climate and therefore not entirely obvious a driver of its change.

Most calculations in peer-reviewed literature (or not) that focus on the CO_2 factor indeed would only like us to believe that CO_2, especially under the realm of radiative forcing, is the predominant factor for driving *anomalous* climate responses, while the unavoidable and very difficult core subject about the actual dynamical state of Earth's "mean" climate is ignored.

Third, some 10 years ago, Lindzen (1994, Annual Review in Fluid Mechanics, vol. 26, 353–378) pointed out a rather serious internal inconsistency regarding the role of water vapor and clouds when the physics of greenhouse effect is normally evaluated even among expert scientists or expert sources of information. (See e.g., the comment "without [the greenhouse effect], the planet would be 65 degrees colder" by Jerry Mahlman in the February 2004 issue of Crisis Magazine, *http://www.crisismagazine.com/february2004/feature1.htm*) and the description of Greenhouse Effect in the EPA's "global warming for kids" webpage: *http://www.epa.gov/globalwarming/kids/greenhouse.html.*) Lindzen notes the "artificial inevitability" for the predominance of CO_2 radiative forcing as a climatic factor in the following passage.

> "In most popular depictions of the greenhouse effect, it is noted that in the absence of greenhouse gases, the Earth's mean temperature would be 255 K [about 0° F], and that the presence of infrared absorbing gases elevates this to 288 K [59° F]. In order to illustrate this, only radiative heat transfer is included in the schematic illustrations of the effect (Houghton *et al.* 1990, 1992) [IPCC reports]; this lends an artificial inevitability to the picture. Several points should be made concerning this picture: 1. The most important greenhouse gas is water vapor, and the next most important greenhouse substance consists in clouds; CO_2 is a distant third (Goody & Yung 1989). 2. In considering an atmosphere without greenhouse substances (in order to get 255 K), clouds are retained for their visible reflectivity while ignored for their infrared properties. More logically, one might assume that the elimination of water would also lead to the absence of clouds, leading to a temperature of about 274 K [or 278 K depending on what value of the solar irradiation factor is used] rather than 255 K. 3. Pure radiative heat transfer leads to a surface temperature of about 350 K rather than 288 K. The latter temperature is only achieved by including a convective adjustment that consists simply in adjusting vertical temperature gradient so as to avoid convective instability while maintaining a consistent radiative heat flux. . . . " (p. 359–361)[2]

Hu *et al.* (2000, *Geophysical Research Letters*, vol. 27, 3513–3516) added that as the sophistication of parameterization of atmospheric convection increases, there is a tendency for climate model sensitivity to variation in atmospheric CO_2 concentration to decrease considerably. In Hu *et al.* (2000)'s study, the change is from a decrease in the averaged tropical warming of 3.3 to 1.6° C for a doubling of CO_2 that is primarily associated the corresponding decrease in the calculated total atmospheric column increase in water vapor from 29 percent to 14 percent.

Question 26. If you accept those radiation calculations as valid, please explain why you seem to believe that the energy trapped by the greenhouse gases will have a small effect whereas you seem to believe that small changes in solar energy will have very large climatic effects?

Response. In addition to my answers under Questions No. 19, 20 and 25 above, I would like to point out that the Sun's radiation is not only variable but it varies in the ultraviolet part of the electromagnetic spectrum often by factors of 10 or more. The question about the relative effects of anthropogenic greenhouse gases and the Sun's radiation in terms of radiative forcing is certainly of interest but it does not add much to my current research quest to understand the Earth's mean climatic state and its nonlinear manifestations.

Question 27. Please explain why you think the physically based climate models seem to quite satisfactorily represent the seasonal cycles of the climate at various latitudes based on the varying distributions of solar and infrared energy, but then would be so far off in calculating the climatic response for much smaller perturbations to solar radiation and greenhouse gases?

Response. As indicated below, the first part of this sentence about a satisfactory representation of seasonal cycles of climate by computer climate models is not any assured statement of fact. This is why the followup question cannot be logically answered.

---

[2] A more pedagogical discussion of the greenhouse effect is given by Lindzen and Emanuel (2002) in *Encyclopedia of Global Change, Environmental Change and Human Society*, Volume 1, Andrew S. Goudie, editor in chief, p. 562–566, Oxford University Press, New York, 710 pp.

For example, E. K. Schneider (2002, *Journal of Climate*, vol. 15, 449–469) noted that:

> "[a]t this writing, physically consistent and even flux-corrected coupled atmosphere-ocean general circulation models (CGCMs) have difficulty in producing a realistic simulation of the equatorial Pacific SST [sea surface temperature], including annual mean, annual cycle, and interannual variability. Not only do the CGCM simulations have significant errors, but also there is little agreement among models."

In a systematic comparison of the performance of 23 dynamical ocean-atmosphere models, Davey *et al.* (2002, Climate Dynamics, vol. 18, 403–420) found that "no single model is consistent with the observed behavior in the tropical ocean regions . . . as the model biases are large and gross errors are readily apparent." Without flux adjustment, most models produced annual mean equatorial sea surface temperature in the central Pacific that are too cold by 2–3° C. All GCMs except one simulated the wrong sign of the east-west SST gradient in the equatorial Atlantic. The GCMs also incorrectly simulated the seasonal climatology in all ocean sections and its interannual variability in the Pacific ocean.

Question 28. In regard to your answers to the previous questions, to what extent is your indication of a larger climate sensitivity for solar than greenhouse gases due to quantitative analysis of the physics and to what extent due to your analysis of statistical correlations? Is this greater responsiveness for solar evident in the baseline climate system, or just for perturbations, and could you please explain?

Response. Please see my answers to Questions No. 26 above and 30 below.

Question 29. Please explain why you seem to accept that solar variations, volcanic eruptions, land cover change, and perhaps other forcings can have a significant climatic influence, but changes in CO_2 do not or cannot have a comparable influence?

Response. Please see my answers to Question No. 30.

Question 30. Could you please clarify why it is that you think the best way to get an indication of how much the climate will change due to global-scale changes in greenhouse gases or in solar radiation is to look at the regional level rather than the global scale? How would you propose to distinguish a natural variation from a climate change at the local to regional level?

Response. Questions No. 28, 29 and 30 seem to be based on the unreasonable presumptions that some special insights about the effects of solar irradiation or land cover changes or even volcanic eruptions must be invoked or answered in order to challenge the role of carbon dioxide forcing in the climate system. That presumption is illogical. My basic view and research interest about carbon dioxide and the ongoing search for the right tool for modeling aspects of the Earth's climate system can be briefly summarized by my answers to Questions No. 19, 25, 26, 27 and perhaps 20.

As to your specific question on distinguishing a natural variation (either internally generated or externally introduced by solar variation or volcanic eruption) from a climate change by anthropogenic factors like land cover changes or carbon dioxide at the local to regional level, there is possibly a somewhat surprising answer. If one wish to single out the potential effects of man-made carbon dioxide against other natural and anthropogenic factors as hinted by your question, then the answer is clear—the CO_2 effect is expected to be small in the sense that its potential signals will be likely be overwhelmed when compared with expected effects by other factors. It is a scientific fact that the signal of CO_2 on the climate may be expected only over a very long time baseline and over a rather large areal extent. For example, Zhao and Dirmeyer (2003, COLA Technical Report No. 150; available at *http://grads.iges.org/pubs/tech.html*), in their modeling experiments that attempt to account for the realistic effects of land cover changes, sea surface temperature changes and for the role of added atmospheric CO_2, found that

> "[w]hen observed CO_2 concentrations are specified in the model across the 18-year period, . . . we do not find a substantially larger warming trend than in CTL [with no change in CO_2 concentration], although some small increase is found. The weak impact of atmospheric CO_2 changes may be due to the small changes in specified CO_2 during the model simulation compared to the doubling CO_2 simulation, or the short length of the integrations. It is clear that the relatively strong SST [sea surface temperature] influence in this climate model is the driver of the [observed] warming."

Please also consider the point made by Lindzen (2002) under Question No. 8 above concerning the difficulties in linking the observed warming trend of the deep ocean (without challenging the quality and error of those deep ocean temperature data)

to anthropogenic CO_2 forcing. Finally, I wish to note that Mickley *et al.* (2004, *Journal of Geophysical Research*, vol. 109, D05106) managed to use climate model simulations results to demonstrate "the limitations in the use of radiative forcing as a measure of relative importance of greenhouse gases to climate change. . . . While on a global scale CO_2 appears to be a more effective 'global warmer' than tropospheric ozone per unit forcing, regional sensitivities to increase ozone may lead to strong climate responses on a regional scale."

Question 31. How does your recent article relate to your assignments at the Harvard Smithsonian Observatory? Is paleoclimate part of the task of this observatory?

Response. The publications of Soon *et al.* (2003) or Soon *et al.* (2004) are possible because of research grants that I and my collaborators obtained through competitive proposals to several research funding sources. I am a trust-fund employee at the Harvard-Smithsonian Center for Astrophysics and the support of my position and research work here is mainly through my own research initiative and proposal application. The scientific learning about paleoclimatic reconstruction presented in Soon *et al.* (2003) is related to my research interest in the mechanisms of sun-climate relation, especially for relevant physical pathways and processes on multidecadal and centennial time scales. Additional fruit of my independent research and labor in the area of sun-climate physics, funded or unfunded, is exemplified by the March 2004 book "The Maunder Minimum and The Variable Sun-Earth Connection" (see *http://www.wspc.com/books/physics/5199.html*) by W. Soon and S. Yaskell (published by World Scientific Publishing Company). It might also be instructive to note that paleoclimate researchers have been speculating about long-term variability of the sun as the cause of centennial- to millennial-scale variability seen in their proxy records.

Question 32. In your testimony, you said that "climate change is part of nature." Please describe what you meant, since obviously, climate change have occurred due, in part, to changes in various forcings, such as solar, continental drift, atmospheric composition, asteroid impacts, etc. rather than being just completely random events. Could you provide estimates of how large you consider future forcings might be and how big the climate change they might cause could be?

Response. In this occasion, I am referring to the fact that any change or variability in climate is most likely a rule, rather than the exception, of the climate system. But I was not speaking about or trying to imply the factors of change, either naturally produced or man-made. I apologize for any potential confusion. It is certainly reasonable to suggest that those climatic changes may arise from "forcings" but it would be unwise to rule out internally generated manifestations of climatic variables that could be purely stochastic in origin. I would strongly recommend the pedagogical discussion by Professor Carl Wunsch of MIT in Wunsch (1992, Oceanography, vol. 5, 99–106) and Wunsch (2004, "Quantitative estimate of the Milankovitch-forced contribution to observed Quaternary climate change", working manuscript downloadable from *http://puddle.mit.edu/*~cwunsch/).

I cannot speculate on future climate forcings and resultant climatic changes because I found no basis for doing so.

Question 33. Please provide a comparable estimate, with some supporting examples from the past, of how big you think the decadal (or 50-year if you prefer) change in the hemispheric/global climate could be due to natural variability? If you prefer to focus on the regional scale change, could you provide an indication of any expected change in the degree of regional variability about the hemispheric and global values, and what the mechanism for this might be?

Response. This question seems a related question trying to get at a quantitative comparison of how large natural climate variability on regional or hemispheric scale can be under the shadow of expected future changes. Again, with no intention to devalue this interesting question, I do not have sufficient knowledge nor ability to venture such an estimate. In fact, I would go so far to say that if the estimates of variability for both the past and future are known within a reasonable range of uncertainties, then the actual scientific research program to address questions about the role of added carbon dioxide no longer require further funding or execution since we have obtained all the relevant answers. But you may have judged from my answers given throughout this Q&A that much remains to be quantified and understood and the hard scientific research must continue.

Question 34. Please explain the scientific basis for your testimony that "one should expect the CO_2 greenhouse effect to work its way downward toward the surface."

Response. Co-answer to this question is given under Question No. 35.

Question 35. Do you believe that there is greater greenhouse trapping of energy in the troposphere than at the surface and that the atmosphere has a low heat capacity? If so, how big is this temperature difference?

Response. It is broadly agreed and assumed that carbon dioxide, when released into the air, has a tendency to get mixed up quickly and so is distributed widely through out the whole column of the atmosphere. The air near the surface is already dense and moist, so addition of more carbon dioxide will introduce very little imbalance of radiation energy budget there. In contrast adding more carbon dioxide to the thinner and drier air of the troposphere will cause a chain of noticeable effects. First, the presence of more carbon dioxide in the uppermost part of the atmosphere will cause more infrared radiation to escape into space because there are more carbon dioxide molecules to channel this infrared radiation upward and outward unhindered. Part of that infrared radiation is also being emitted downward to the lower parts of the atmosphere and the surface where it is reabsorbed by carbon dioxide and the thicker air there. The layer of air at the lower and middle troposphere, being more in direct contact with this down-welling radiation, is expected to heat more than air near the surface. Thus, adding more carbon dioxide to the atmosphere should cause more warming of the air around the height of two to seven kilometers. (Please consider for example the discussion by Kaser *et al.* (2004) under Question No. 8 about the ineffectiveness of an added longwave radiation from a direct addition of atmospheric CO_2 or atmospheric temperature change in explaining the modern retreat of glaciers at Kilimanjaro.) In other words, the clearest impact of the carbon dioxide greenhouse effect should manifest itself in the lower- and mid-troposphere rather than near the earth's surface. Here, I am mostly speaking on the basis of expectation from pure radiative forcing considerations.

Such a qualitative description is not complete, even though that is roughly what was modeled in the most sophisticated general circulation models (see e.g., Chase *et al.*, 2004, Climate Research, vol. 25, 185–190), because it misses the key roles of atmospheric convection and waves as well as all the important hydrologic processes (please see e.g., Neelin *et al.*, 2003, *Geophysical Research Letters*, vol. 30 (no. 24), 2275 and consider additional remarks about water vapor and atmospheric convection under Question No. 25 as well as discussion on climate forcing factors and climate response sensitivities under Question No. 20). Some theoretical proposals expect a warming of the surface relative to the low- and mid-troposphere because of nonlinear climate dynamics (Corti *et al.*, 1999, Nature, 398, 799–802). That expectation is because of the differential surface response with the pattern of Cold Ocean and Warm Land (COWL) that becomes increasingly unimportant with distance away from the surface (rather than just the difference in heat capacity mentioned in your question) [see Soon *et al.*, 2001 for additional discussion]. Nevertheless, no GCM has yet incorporated such an idea into an operationally robust simulation of the climate system response to greenhouse effects from added CO_2. In the latest "global warming" work, Neelin *et al.* (2003), for example, still distinctly differentiate between mechanisms for tropical precipitation that are initiated through CO_2 warming of the troposphere and through El Niño warming rooted in oceanic surface temperature and subsurface thermocline dynamics. (Further note that their model experiments [see Figure 2b+2c and 10b+10c of Chou and Neelin, 2004, "Mechanisms of global warming impacts on regional tropical precipitation" in preparation for *Journal of Climate*; available at ***http://www.atmos.ucla.edu/?csi/REF/***] also clearly shown that the troposphere warmed significantly more than surface with the doubling of atmospheric CO_2 as discussed by Chase *et al.* 2004 below.)

But it is worth noticing that the current global observation shows that, at least over the 1979–2003 interval, the lower tropospheric temperatures are not warming as fast as the surface temperatures (see Christy *et al.* 2003, Journal of Atmospheric and Oceanic Technology, vol. 20, 613–629; for additional confidence on the results derived by the University of Alabama-Huntsville group, please see Christy and Norris, 2004, *Geophysical Research Letters*, vol. 31, L06211). This observed fact is in contradiction to the accelerated warming of the mid and upper troposphere relative to surface simulated in current models (Chase *et al.* 2004). Chase *et al.* (2004) arrives at the following conclusions, upon examining results from 4 climate models in both unforced scenarios and scenarios forced with increased atmospheric greenhouse gases and the direct aerosol effect[3]:

[3] Such a study should also be consistently challenged by the discussion under question No. 20 about the adequacy of studying responses from a combination of incomp;ete forcings—through my primary purpose here is to illustrate the theoretical expectation of CO_2 forcing deriving from state-of-the-art climate models.

- "Model simulations of the period representative of the greenhouse-gas and aerosol forcing for 1979–2000 generally show a greatly accelerated and detectable warming at 500 mb relative to the surface (a 0.06° C decade^{-1}increase).
- Considering all possible simulated 22 yr trends under anthropogenic forcing, a strong surface warming was highly likely to be accompanied by accelerated warming at 500 mb [i.e., 987 out of 1018 periods or 97 percent of the cases had a larger warming at 500 mb than at the surface] with no change in likelihood as forcings increased over time.
- In simulated periods where the surface warmed more quickly than 500 mb, there was never a case [emphasis added] in which the 500 mb temperature did not also warm at a large fraction of the surface warming. A 30 percent acceleration at the surface was the maximum simulated as compared with an observed acceleration factor of at least 400 percent the mid-troposphere trend.
- In cases where there was a strong surface warming and the surface warmed more quickly than at 500 mb in the forced experiments, there was never a case in which the 500 mb-level temperatures did not register a statistically significant ($p< 0.1$) trend (i.e., a trend detectable with a simple linear regression model). The minimum p value of approximately 0.08 occurred in the single case in which the significance was not greater than 99 percent.
- It was more likely that surface warmed relative to the mid-troposphere under control simulations than under forced simulations.
- At no time, in any model realization, forced or unforced, did any model simulate the presently observed situation of a large and highly significant surface warming accompanied with no warming whatsoever aloft." (p. 189)

Question 36. The grants that are described as supporting your analysis seem to have much more to do with the sun or unrelated pattern recognition that with climate history (Air Force Office of Scientific Research-Grant AF49620–02–1–0194; American Petroleum Institute-Grants 01–0000–4579 and 2002–100413; NASA-Grant NAG–7635; and NOAA-Grant NA96GP0448). Could you please describe how much funding you received and used in support of this study, all of the sources and the duration of that funding, and the relevance of those grant topics to the article?

Response. All sources of funding for my and my colleagues' research efforts that resulted in the publication of Soon and Baliunas (2003) and Soon *et al.* (2003) were openly acknowledged. In other words, all sources of funding were disclosed in the manuscripts when they were submitted for publication; all sources of funding were also disclosed to readers in the printed journal articles. I am not the principal investigator for some of the grants we received (e.g., the NOAA grant was awarded to Professor David Legates), so I am not in the privilege position to provide exact quantitative numbers. But throughout the 2001–2003 research interval in which our work was carried out, the funding we received from the American Petroleum Institute was a small fraction of the funding we received from governmental research grants.

The primary theme of my research interest is on physical mechanisms of the sun-climate relationship. This is why researching into the detailed patterns of local and regional climate variability as published in Soon *et al.* (2003) is directly relevant to that goal. Please also consider my research position listed under Question No. 31 above.

Question 37. Have you been hired by or employed by or received grants from organizations that have taken advocacy positions with respect to the Kyoto Protocol, the U.N. Framework Convention on Climate Change, or legislation before the U.S. Congress that would affect greenhouse gas emissions? If so, please identify those organizations.

Response. I have not knowingly been hired by, nor employed by, nor received grants from any such organizations described in this question.

Question 38. Please describe the peer review process that took place with respect to your nearly identical articles published both in Climate Research and in Energy and Environment, including the number of reviewers and the general content of the reviewers' suggested edits, criticisms or improvements.

Response. The Climate Research paper (Soon and Baliunas, 2003, Climate Research, vol. 23, 89–110) was submitted for publication and went through a routine peer-review process and was eventually approved for publication. The main content of the review was to propose: (a) reorganizing of materials including elimination of discussions on ENSO and GCMs; (b) removing "tone" problems by eliminating criticisms of previous EOF and superposition analyses; (c) reducing quotes especially those by Hubert Horace Lamb to improve readability; and (d) reviewing changes in each region with same thoroughness. The July 3, 2003's email (as Attachment I

below) from the director of Inter-Research, Otto Kinne, who publishes Climate Research is enclosed below to confirm that the review process was fairly rigorous and all parties involved had carried out their roles and duties in this time-honored system properly.

The extended and more complete paper by Soon *et al.* (2003, Energy & Environment, vol. 14, 233–296) was submitted to Energy & Environment for consideration together with the accepted Climate Research manuscript. Energy & Environment's editorial decision was to send our manuscript for review, and after acceptance, include in its editorial in Energy & Environment, volume 14, issues 2&3, a footnote referring to the Climate Research paper.

Finally, we wish to correct that the false impression introduced by Professor Mann both during the testimony and in public media that his attack on the papers by Soon and Baliunas (2003) and Soon *et al.* (2003), in a FORUM article in the American Geophysical Union *Eos* newspaper (Mann *et al.*, 2003, *Eos*, vol. 84, 256–258), were either rigorously peer-reviewed or represented widespread view of the community. Contrary to Professor Mann's public statements, a FORUM article in *Eos* is said to be only stating "a personal point of view" (*http://www.agu.org/pubs/eos—guidelines.html#authors*). Whatever peer-reviewing that was done did not include soliciting comments from the authors of the papers being criticized. We first learned of this FORUM article from the AGU's press release No. 03–19 "Leading Climate Scientists Reaffirm View that Late 20th Century Warming Was Unusual and Resulted From Human Activity" (*http://www.agu.org/sci—soc/prrl/prrl0319.html*). See Soon *et al.* (2003b, *Eos*, vol. 84 (44), 473–476) for our own response to the Mann *et al.* FORUM article.

Statement of Professor Michael E. Mann, Department of Environmental Services, University of Virginia

My name is Michael Mann. I am a professor in the Department of Environmental Sciences at the University of Virginia. My research involves the use of climate models, the analysis of empirical climate data, and statistical methods for comparing observations and model predictions. One area of active current research of mine involves the analysis of climate "proxy" records (that is, natural archives of information which record past climate conditions by their biological, physical, or chemical nature). These data are used to reconstruct patterns of climate variability prior to the period of the past century or so during which widespread instrumental climate records are available. A primary focus of this research is deducing the long-term behavior of the climate system and the roles of various potential agents of climate change, both natural and human.

I was a Lead Author of the "Observed Climate Variability and Change" chapter of the *Intergovernmental Panel on Climate Change* (IPCC) Third Scientific Assessment Report and a scientific contributor for several other chapters of the report. I am the current organizing committee chair for the National Academy of Sciences '*Frontiers of Science*' and have served as a committee member or advisor for other National Academy of Sciences panels related to climate change. I have served as editor for the '*Journal of Climate*' of the American Meteorological Society. I'm a member of the advisory panel for the National Oceanographic and Atmospheric Administrations' *Climate Change Data and Detection Program*, and a member of numerous other international and U.S. scientific working groups, panels and steering committees. I have co-authored more than 60 peer-reviewed articles and book chapters on diverse topics within the fields of climatology and paleoclimatology. Honors I have received include selection in 2002 as one of the 50 leading visionaries in Science and Technology by *Scientific American* magazine, the outstanding scientific publication award for 2000 from the National Oceanographic and Atmospheric Administration, and citation by the Institute for Scientific Information (ISI) for notable recognition of my peer-reviewed research by fellow scientists.

In my testimony here today, I will explain: (1) How mainstream climate researchers have come to the conclusion that late-20th century warmth is unprecedented in a very long-term context, and that this warmth is likely related to the activity of human beings. (2) Why a pair of recent articles challenging these conclusions by astronomer Willie Soon and his co-authors are fundamentally unsound.

CLIMATE HISTORY AND ITS IMPLICATIONS

Evidence from paleoclimatic sources overwhelmingly supports the conclusion that late-20th century hemispheric-scale warmth was unprecedented over at least the past millennium and probably the past two millennia or longer.

Modeling and statistical studies indicate that such anomalous warmth cannot be explained by natural factors but, instead, requires significant anthropogenic (that is, 'human') influences during the 20th century. Such a conclusion is the indisputable consensus of the community of scientists actively involved in the research of climate variability and its causes. This conclusion is embraced by the position statement on "Climate Change and Greenhouse Gases" of the American Geophysical Union (AGU) which states that there is a compelling basis for concern over future climate changes, including increases in global-mean surface temperatures, due to increased concentrations of greenhouse gases, primarily from fossil-fuel burning. This is also the conclusions of the 2001 report of the Intergovernmental Panel on Climate Change (IPCC), affirmed by a National Academy of Sciences report solicited by the Bush administration in 2001 which stated, *"The IPCC's conclusion that most of the observed warming of the last 50 years is likely to have been due to the increase in greenhouse gas concentrations accurately reflects the current thinking of the scientific community on this issue."*

THE MAINSTREAM SCIENTIFIC VIEWPOINT

Human beings have influenced modern climate through changes in greenhouse gas concentrations, the production of industrial aerosols, and altered patterns of land-use. By studying both the record of ancient climate variability and the factors that may have influenced it, we can establish how and why the climate system varied naturally, prior to any large-scale anthropogenic impacts. Large changes in climate certainly occurred in the distant past. If we look 60 million years back in time, Dinosaurs were roaming the polar regions of the earth, and the globe was almost certainly warmer than today. Carbon dioxide levels were probably about double their current level, and had slowly attained such high levels due to changes in the arrangements of the continents ('plate tectonics') which influence the outgassing of carbon dioxide from the solid earth and thus, atmospheric greenhouse gas concentrations. These changes occur on timescales of tens of millions of years. 10,000 years ago, large ice sheets existed over North America due to natural changes that occur in the earth's orbit on timescales of tens of thousands of years. Trying to study distant past climates for insights into modern natural climate variability is hampered by the fact that the basic external constraints on the system (the continental arrangement, the geometry of the earth's astronomical orbit, the presence of continental ice sheets—what we call the 'boundary conditions') were significantly different from today. Focusing on the evolution of climate in *the centuries* leading up to the 20th century provides a perspective on the natural variability of the climate prior to the period during which large-scale human influence is likely to have occurred, yet modern enough that the basic boundary conditions on the climate system were otherwise the same. This provides us, in essence, a 'control' for diagnosing whether or not recent climate changes are indeed unusual.

Instrumental data for use in computing global mean surface temperatures are only available for about the past 150 years. Estimates of surface temperature changes prior to the 20th century must make use of historical documents and natural archives or "proxy" indicators, such as tree rings, corals, ice cores and lake sediments, to reconstruct the patterns of past climate change. Due to the paucity of data in the Southern Hemisphere, recent studies have emphasized the reconstruction of Northern Hemisphere rather than global mean temperatures. A number of independent reconstructions of the average temperature of the Northern Hemisphere support the conclusion that the hemispheric warmth of the late 20th century (i.e., the past few decades) is likely unprecedented over at least the past millennium. Preliminary evidence suggests that such a conclusion may well hold for at least the past two millennia, though more work, requiring the development of a more complete set of reliable proxy records spanning the past few millennia, are necessary to further decrease the uncertainties. Climate model simulations employing estimates of natural and anthropogenic radiative forcing changes agree well with the proxy-based reconstructions (Figure 1). The simulations, moreover, show that it is not possible to explain the anomalous late 20th century warmth without the contribution from anthropogenic influences. Such consensus findings are expressed in the recently published article co-authored by myself and 12 other leading climate scientists from the United States and Britain that appeared recently in the journal '*Eos*', the official transactions of the American Geophysical Union, the largest professional society in the field.

FLAWS IN A RECENT STUDY DISPUTING THE SCIENTIFIC CONCENSUS

Two deeply flawed (and nearly identical) recent papers by astronomers Soon and Baliunas (one of them with some additional co-authors—both henceforth referred to

as 'SB') have been used to challenge the scientific consensus. I outline the 3 most basic problems with their papers here:

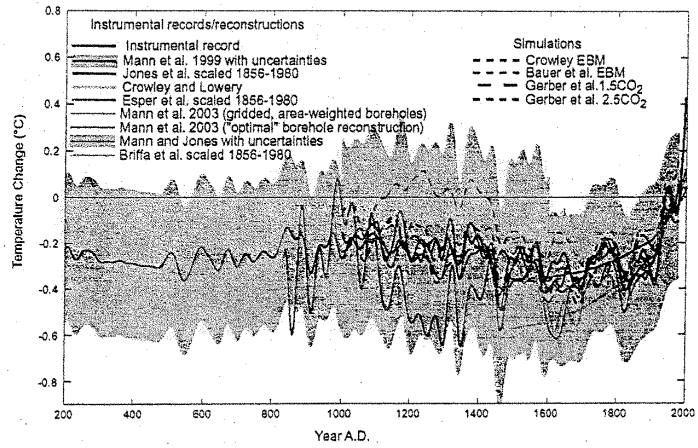


Figure 1. Estimates of changes in the average temperature of the northern hemisphere over the past one to two millennia based on reconstructions from empirical ('proxy') data, and from models driven with estimated changes in natural and anthropogenic factors. Also shown is the modern (1856-2002) instrumental record. Values shown are relative temperatures (in °C) from the 1961-1990 average (from: Mann, M.E., Ammann, C.M., Bradley, R.S., Briffa, K.R., Crowley, T.J., Hughes, M.K., Jones, P.D., Oppenheimer, M., Osborn, T.J., Overpeck, J. T., Rutherford, S., Trenberth, K.E., Wigley, T.M.L., On Past Temperatures and Anomalous Late 20th Century Warmth, *Eos*, 84, 256-258, 2003.)

(1) In drawing conclusions regarding past regional temperature changes from proxy records, it is essential to assess to make sure that the proxy data are indicators of *temperature* and not precipitation or drought. SB make this fundamental error when they cite evidence of either 'warm', 'wet', or 'dry' regional conditions as being in support of an exceptional 'Medieval Warm Period' or 'MWP'. Their criterion, *ad absurdum*, could be used to define *any* period of climate as 'warm' or 'cold'. *Experienced* paleoclimate researchers know that they must first establish the existence of a temperature signal in a proxy record before using it to evaluate past temperature changes (Figure A1).

(2) It is essential to distinguish between regional temperature changes and truly hemispheric or global changes. SB do not make this essential distinction. The wavelike character of weather (i.e., the day-to-day wiggles of the Jet Stream) ensures that certain regions tend to warm when other regions cool. This past winter is a case in point. January was about 2° C *below* normal on the east coast of the U.S., but about 4° C *above* normal over much of the west. Utah, Nevada and parts of California and Alaska had the warmest January on record (the change in location of the Iditarod dog sled race was a casualty of the Alaskan winter warmth!). The average temperature over the entire U.S. was about 1° C above normal, much less warm than for the western U.S., and of the opposite sign of the eastern U.S.

In a similar manner, *average global or hemispheric* temperature variations on longer timescales tend to be much smaller in magnitude than those for particular regions, due to the tendency for a cancellation of warm and cold conditions in different regions. While relative warmth during the 10th–12th centuries, and cool conditions during the 15th-early 20th centuries are evident from reconstructions and model simulations of the average temperature of the Northern Hemisphere (Figure 1), the specific periods of cold and warmth naturally differ from region to region (Figure A2). The notion of an unusually cold 17th century 'Little Ice Age', for example, arose in a European historical context. What makes the late 20th century unique is the simultaneous warmth indicated by nearly all long-term records (Figure A2), leading to the anomalous warmth evident during this period in Northern Hemisphere average temperatures (Figure 1).

(3) It is essential, in forming a climate reconstruction, to carefully define a base period for modern conditions against which past conditions may be quantitatively compared. The concensus conclusion that late–20th century mean warmth likely exceeds that of any time during the past millennium for the Northern Hemisphere,

is based on a careful comparison of temperatures during the most recent decades with reconstructions of past temperatures, *taking into account the uncertainties in those reconstructions*. As it is only the past few decades during which Northern Hemisphere temperatures have exceeded the bounds of natural variability, any analysis such as SB that compares past temperatures only to early or mid-20th century conditions, or interprets past temperatures using proxy information not capable of resolving decadal trends cannot address the issue of whether or not late-20th century warmth is anomalous in a long-term context.

CONCLUSIONS

The concentration of greenhouse gases in the atmosphere is higher than at any time in *at least* the last 400,000 years, and, it increasingly now appears, probably many millions of years. This increase is undeniably due to the activity of human beings through fossil fuel burning. Late 20th century warming is unprecedented in modern climate history at hemispheric scales. This is almost certainly a result of the dramatic increase in greenhouse gas concentrations due human activity. The latest model-based projections indicate a global mean temperature increase of 0.6 to 2.2° C (1° C to 4° F) relative to 1990 levels by the mid-20th century. While these estimates are uncertain, even the lower value would take us well beyond any previous levels of warmth seen over at least the past couple millennia. The magnitude of warmth, but perhaps more importantly, *the unprecedented rate of this warming*, is cause for concern.

Appendix 1.

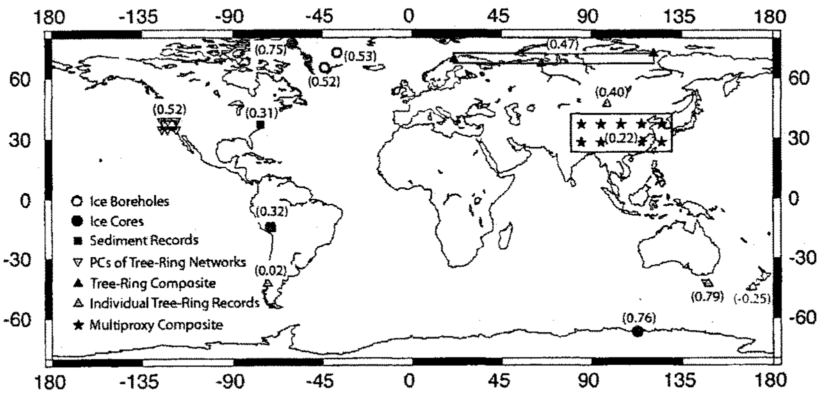


Figure A1. Locations of temperature proxy records used to reconstruct past hemispheric temperature trends (from: M.E. Mann, and P.D. Jones, 2003: *Geophysical Research Letters*, v.26, p.759-762).

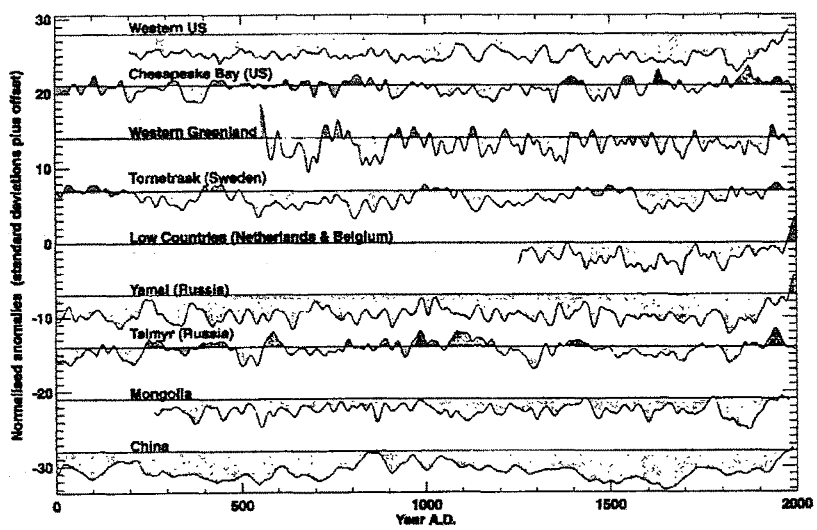


Figure A2. Temporal histories of nine temperature-sensitive proxy records across the Northern Hemisphere, smoothed to highlight changes on 40 year and longer timescales. Blue (red) shading indicates values below (above) the 1961–1990 means (from: Mann, M.E., Ammann, C.M., Bradley, R.S., Briffa, K.R., Crowley, T.J., Hughes, M.K., Jones, P.D., Oppenheimer, M., Osborn, T.J., Overpeck, J. T., Rutherford, S., Trenberth, K.E., Wigley, T.M.L., On Past Temperatures and Anomalous Late 20th Century Warmth, *Eos*, 84, 256-258, 2003.)

RESPONSE BY MICHAEL MANN TO ADDITIONAL QUESTIONS BY SENATOR INHOFE

Question 1. You have used the term "climate scientist" to distinguish certain individuals. What, in your view, does it take for one to earn the title "climate scientist"? What specific credentials, or the lack thereof, would lead you to refuse to recognize someone as a "climates scientist"?

Response. The term "climate scientist" is used, in my experience, to describe an individual with specific training in oceanographic, atmospheric, and coupled ocean-atmosphere processes relevant to understanding climate variability and the behavior of the climate system. An individual might obtain this training through either an advanced degree in those areas of study, or through years of research in those areas associated with numerous publications in the peer-reviewed climate literature such as "*Journal of Geophysical Research—Atmospheres*", "*Journal of Geophysical Research-Oceans*", "*Climate Dynamics*", "*The Holocene*", "*Geophysical Research Letters*", "*Paleoceanography*" (or publication of climate papers in leading international science journals such as "*Nature*" and "*Science*"). I would not, for example, consider scientists with advanced degrees in Astronomy, Astrophysics, or Physics who have published primarily in those areas, as "climate scientists"—nor do I believe would most of my colleagues in the climate research community. In addition to training and publishing in a field, leading scientists would normally be expected to be actively interacting and collaborating in studies with colleagues and ensuring their understanding of cutting edge science through attendance and active participation in meetings convened by the leading professional societies and organizations.

Question 2. Your work and testimony contends that the Little Ice Age was not global, but restricted to only portions of Europe. A forthcoming article by Shindell *et al.* (Shindell, D.T. *et al.*, 2003: Volcanic and solar forcing of climate change during the pre-industrial era. *Journal of Climate*, in press), however, indicates the Little Ice Age could have resulted from a combination of solar and volcanic forcing. Do you agree with these conclusions from Shindell *et al.*? If so, how can solar and volcanic forcings generate climatic effects that are not observed across the entire hemisphere?

Response. The statement is incorrect. I never testified that the "Little Ice Age was . . . restricted to only portions of Europe".

It should first be noted that many paleoclimatologists have questioned the utility of terms such as "Little Ice Age" and "Medieval Warm Period" which provide misleading descriptions of past climate changes in many regions. There is a complex pattern of climate variability in past centuries, and the lack of evidence for synchronous temperature variations worldwide in past centuries [e.g. Bradley, R.S., and P.D. Jones, "Little Ice Age" summer temperature variations: their nature and relevance to recent global warming trends, *The Holocene*, 3, 367–376, 1993; Hughes, M.K., and H.F. Diaz, Was there a 'medieval warm period', and if so, where and when, *Climatic Change*, 26, 109–142, 1994]. The cited paper by Shindell et al (2003), of which I am a co-author, is fully consistent with such findings. The paper, rather than demonstrating globally uniform patterns of warming or cooling in past centuries, shows that surface temperature changes were dominated by regional overprints associated with the response of the "North Atlantic Oscillation" atmospheric circulation pattern to radiative forcing. This response leads to a pattern of cooling during the 17th/18th centuries in certain regions (not just Europe, but many regions throughout the Northern Hemisphere extratropics) and warming in other regions. The paper shows that this pattern of warming and cooling closely resembles the pattern of surface temperature change during that interval reconstructed by Mann and colleagues (MBH98). It is worth noting, moreover, that the tropical Pacific seems to have been in a warmer, rather than a "colder" state, during the conventionally defined "Little Ice Age" [Cobb, K.M., Charles, C.D., Edwards, R.L., Cheng, H., & Kastner, M. El Niño-Southern Oscillation and tropical Pacific climate during the last millennium, Nature 424, 271–276 (2003). Climate dynamists understand the importance of such phenomena in understanding the highly variable pattern of surface temperature changes in past centuries, and rarely, if ever, argue for the existence of globally uniform or synchronous temperature change in past centuries. The response of the climate to solar and volcanic radiative forcing is known to involves dynamical responses associated with regionally differentiated temperature trends that overprint far smaller global mean responses. This contrasts strongly with the response of the climate to anthropogenic climate forcing, for which the integrated global mean radiative forcing is considerably greater, and the associated large-scale warming typically rises above the regional variability.

Question 3. That same paper finds "long-term regional response to solar forcing [that] greatly exceeds unforced variability . . . and produces climateanomalies simi-

lar to those seen during the Little Ice Age. Thus, longterm regional changes during the pre-industrial [era] appear to havebeen dominated by solar forcing.'' You further state that ''For the few centuries prior to the industrial era, however, externally driven climate change is thought to have been forced primarily by only two factors: variation in solar output and volcanic eruptions . . . These forcings likely played a large role in the so-called Medieval Warm Period (MWP) and Little Ice Age (LIA) epochs of the last millennium, which saw significant climate changes on at least regional scales . . .'' You then define ''regional'' ''to mean continental in scale . . .'' Do you claim that total solar irradiance change is the only solar forcing mechanism that has any significant climate effect? List your formal training in, plus courses you have taught, in solar physics. Do you agree with the paper's claim that the MWP and LIA exist on regional scales, in accordance with climate experts like R. Bryson and H.H. Lamb, starting with their work in the 1960's, and recently updated in summary in Soon *et al.* (2003)?

Response. Expertise in ''solar physics'' is not the expertise required to evaluate what is happening to the Earth's climate—what matters are the changes in solar radiation at the top of the atmosphere and then down through it. As is made clear in our paper, we are considering not only ''total solar irradiance'' but also its spectral distribution. Indeed, because much of the change in solar radiation occurs in UV wavelengths, induced changes in stratospheric ozone can be lead to significant changes in atmospheric circulation in the troposphere. The model simulations indicate that such atmospheric circulation changes can, acting with other factors, lead to regional variations in the climate such as were observed over the last millennium. As a co-author of this paper, I of course, agree with its findings. However, the inference that this paper confirms the work of Soon *et al.* (2003) is very mistaken.

With respect to my training, teaching, I would encourage that my Curriculum Vitae, which I have provided separately, be included in the record to be compared to those of the other witnesses with respect to relevant expertise and standing in the climate research community.

Question 4. In your testimony, you stated that you hold the ''mainstream'' view with respect to climate theory of air temperature trends over the past two millennia. Provide supporting citations in the refereed scientific literature that are not authored or co-authored by you or your colleagues, collaborators, students or former students, or associates (i.e., Phil Jones, Ray Bradley, Malcolm Hughes), where others hold this Emainstream[1] view.

Response. The statement once again mischaracterizes my comments. As there is only one reconstruction of Northern Hemisphere annual mean temperature over the past two millennia, published only recently by Phil Jones and myself, it is hardly meaningful to discuss whether ''other studies'' support this finding. The peer-review and publication process typically unfolds on timescales of a year or longer, not months. Any careful reading of my comments would reveal that I was not referring to this one specific reconstruction in the comments I made characterizing what I believe to be the mainstream viewpoint of the climate research community. This review, as discussed in my testimony, refers rather to the widespread evidence that late 20th century warmth is unprecedented in a long-term context, anywhere from the past several centuries to nearly the past two millennia, depending on the timeframe of the particular study.

Second, the description of ''collaborators, associates, former students'', depending on how interpreted, is so broad a category as to include just about every leading scientist in the field. The following publications all come to the same conclusion that late 20th century Northern Hemisphere warmth is anomalous in a long-term context:

- Bauer, E., Claussen, M., Brovkin, V., Assessing climate forcings of the earth system for the past millennium, Geophys. Res. Lett., 30 (6), 1276, doi: 10.1029/2002GL016639, 2003.
- Bertrand C., Loutre M.F., Crucifix M., Berger A., Climate of the Last millennium: a sensitivity study. Tellus, 54(A), 221–244, 2002.
- Bradley, R.S., and P.D. Jones, ''Little Ice Age'' summer temperature variations: their nature and relevance to recent global warming trends, The Holocene, 3 (4), 367–376, 1993.
- Bradley, R.S., Briffa, K.R., Crowley, T.J., Hughes, M.K., Jones, P.D, Mann, M.E., Scope of Medieval Warming, Science, 292, 2011–2012, 2001.
- Bradley, R.S., M.K.Hughes and H.F. Diaz., Climate in Medieval Time. Science, 302, 404–405, 2003.
- Pollack, H.N., S. Huang, and P.-Y. Shen, Climate Change Record in Subsurface Temperatures: A Global Perspective, Science, 282, 279–281, 1998.

• Briffa, K.R., T.J. Osborn, F.H. Schweingruber, I.C. Harris, P.D. Jones, S.G. Shiyatov, S.G. and E.A. Vaganov, Low-frequency temperature variations from a northern tree-ring density network. J. Geophys. Res., 106, 2929–2941, 2001.

• Crowley, T.J., Causes of Climate Change Over the Past 1000 Years, Science, 289, 270–277, 2000.

• Crowley, T.J., and T. Lowery, How Warm Was the Medieval Warm Period, Ambio, 29, 51–54, 2000.

• Gerber, S., F. Joos, P. Brügger, T. F. Stocker, M. E. Mann, S. Sitch, and M. Scholze, Constraining temperature variations over the last millennium by comparing simulated and observed atmospheric CO_2, Climate Dynamics, 20, 281–299, 2003.

• Hegerl, G.C., T.J. Crowley, S.K. Baum, K-Y. Kim, and W. T. Hyde, Detection of volcanic, solar and greenhouse gas signals in paleo-reconstructions of Northern Hemispheric temperature. Geophys. Res. Lett., 30 (5), doi: 10.1029/2002GL016635, 2003.

• Huang, S., H. N.Pollack and P.-Y. Shen, Temperature Trends Over the Past Five Centuries Reconstructed from Borehole Temperature, Nature 403, 756–758, 2000.

• Jones, P.D., M. New, D.E. Parker, S. Martin, and I.G. Rigor, 1999: Surface air temperature and its changes over the past 150 years. Reviews of Geophysics 37, 173–199.

• Jones, P.D., T.J. Osborn, and K.B. Briffa, The Evolution of Climate Over the Last Millennium, Science, 292, 662–667, 2001.

• Overpeck, J., K. Hughen, D. Hardy, R. Bradley, R. Case, M. Douglas, B. Finney, K. Gajewski, G. Jacoby, A. Jennings, S. Lamoureux, A. Lasca, G.M.J. Moore, M. Retelle, S. Smith, A. Wolfe, and G. Zielinski, Arctic Environmental Change of the Last Four Centuries, Science, 278, 1251–1256, 1997.

• Pollack, H.N., S. Huang, and P.-Y. Shen, Climate Change Record in Subsurface Temperatures: A Global Perspective, Science, 282, 279–281, 1998.

Question 5. Your work has been characterized as "global" in several venues, including the National Assessment. Is that a fair characterization, or are those sources confused by your use of Northern and Southern Hemisphere proxies in your Northern Hemisphere reconstruction? Can you explain why the National Assessment did not include error bars on your temperature reconstruction?

Response. The proxy records on which our work is based represent conditions over much of the Northern Hemisphere and a small fraction of the Southern Hemisphere. While in any given year there can be some difference in the anomalies in the two hemispheres, the instrumental record indicates that over periods of a few decades or more, the anomalies in the two hemispheres are quite similar because of the thermodynamic and dynamic coupling between them. Thus, the major features of the temperature record, and in particular the unusual 20th century warming, are similar in the two hemispheres and thus global features. It was this aspect of the record to which the text of the National Assessment report refers in presenting the overall significance of our study, and the report is correct in suggesting that the 20th century warming is global in nature. The caption for Figure 2 of Chapter 1 in the Foundation report (page 22 and page 544) states that "Although this record comes mostly from the Northern Hemisphere, it is likely to be a good approximation to the global anomaly based on comparisons of recent patterns of temperature fluctuations." This accurately reflects the situation. In the Overview report (page 13), although the figure title says "Global CO_2 and Temperature Change," the caption next to the figure says "Records of Northern Hemisphere surface temperatures, CO_2 concentrations, and carbon emissions show a close correlation. Temperature change: reconstruction of annual-average Northern Hemisphere surface air temperatures derived from historical records, tree rings, and corals (blue), and air temperatures directly measured (purple)." This quite clearly makes the point this is mainly a Northern Hemisphere temperature record.

With respect to the question about the presentation of the figure, it is misleading to imply that the term "error bars" indicates that the central line is off by this amount—rather the limits mean that there is only 1 chance in 20 that the actual value is outside this range. That is, what we are showing is the likely range within which the anomaly lies, with there being a 95 percent chance the value is within this range. The line that we present in many of our figures, and that was presented in the National Assessment report, is the most likely value within this range (rather a natural choice to display in explaining a complex issue to the public). In looking at the National Assessment report, the caption for Figure 2 of Chapter 1 in the Foundation report (page 22 and page 544) states that "The error bars for the estimate of the annual-average anomaly increase somewhat going back in time, with

one standard deviation being about 0.25° F (0.15° C).” Quite clearly, the reader interested in investigating the accuracy of the records would followup by reading the original reference, which is cited in the text.

In that I was not involved in the National Assessment report, the questioner should consult the authors of that report for any further information or questions.

Question 6. You testified that the late 20th century warming is likely caused byman-made CO_2 forcing on climate; what is your scientific proof for thatclaim Please detail how you removed the potential effects from other factors including those of sulfate aerosols, tropospheric and stratospheric ozone, volcanic dust veiling, black soot, solar particleand wave length-dependent variability, sea ice, land use, vegetation andother greenhouse gases?

Response. The question inconsistently equates my statement of a “likely” causal relationship with the standard of “scientific proof”. Scientists do not speak in terms of “proof”. We speak in terms of likelihoods and the strength of evidence in support of a particular hypothesis.

A large number of peer-reviewed scientific studies have been published in the leading scientific journals such as *Nature and Science* in the past two decades elucidating the role of natural and anthropogenic factors in observed climate changes. Physically based models have been developed and validated against observations, and these models reproduce complex climate phenomena such as El Niño. These same models have been driven with the primary “external” factors that are believed to govern climate variations on timescales of decades and centuries. These external factors include natural factors, such as the modest estimated variations in radiative output of the Sun, which varies by a fraction of a percent over time, variations in the frequency and intensity of explosive volcanic eruptions, which have a several-year cooling effect on the climate through the injection of reflective volcanic aerosols into the stratosphere, and very small changes in the Earth’s orbit relative to the Sun that occur on multi-century timescales. These external factors also include the “anthropogenic” influences of increased greenhouse gas concentrations due to fossil fuel burning, changes in the reflective properties of the land surface due to human land use alterations, and the regional cooling effect of anthropogenic sulphate aerosols in certain industrial regions. When driven with these factors, these climate models have demonstrated a striking ability to reproduce observed global and hemispheric temperature trends during the 20th century, as well as longer-term trends in past centuries as reconstructed from proxy data. Such results have been demonstrated in the following peer-reviewed scientific articles:

- Wigley, T.M.L., R.L. Smith, and B.D. Santer, Anthropogenic Influence on the Autocorrelation Structure of Hemisoheric-Mean Temperatures, Science, 282, 1676–1680, 1998.
- Tett, S.F.B., P.A. Scott, M.R. Allen, W.J. Ingram, and J.F.B. Mitchell, Causes of Twentieth-Century Temperature Change Near the Earth’s Surface, Nature, 399, 569–572, 1999.
- Hegerl, G.C., P.A. Scott, M.R. Allen, J.F.B. Mitchell, S.F.B. Tett, and U. Cubasch, Optimal detection and aftribution of climate change: sensitivity of results to climate model differences, Climate Dynamics, 16, 737–754, 2000.
- Crowley, T.J., Causes of Climate Change Over the Past 1000 Years, Science, 289, 270–277, 2000.
- Stott, P.A., S.F.B. Tett, G.S. Jones, M.R. Allen, J.F.B. Mitchell, and G.J. Jenkins, External Control of 20th Century Temperature by Natural and Anthropogenic Forcings, Science, 290, 2133–2137, 2001.
- Stott, P.A., S.F.B. Tett, G.S. Jones, M.R. Allen, W.J. Ingram, and J.F.B. Mitchell, Attribution of twentieth century temperature change to natural and anthropogenic causes, Climate Dynamics, 17, 11–21, 2001.

These conclusions, furthermore, were endorsed by the 2001 IPCC scientific working group report (Chapter 12), and the followup National Academy of Sciences report that endorsed most of the key IPCC conclusions.

Question 7. A number of expert studies have produced individual proxy records that show the existence of a local Medieval Warm Period or Little Ice Age. Such studies cover a large portion of the globe. How do you reconcile your hemispheric reconstruction with these individual proxy records?

Response. It is unclear to me what precisely the questioner means by “a number of expert studies” or how he defines the “existence” of a “Medieval Warm Period” or “Little Ice Age”. As discussed in my response to question 2, the regionally and temporally variable nature of climate changes in past centuries makes such descriptors of past climate change naive and often useless as a characterization of past changes. A sampling of some of the longest, high-quality best term proxy temperature estimates over the globe was provided in Figure 2 of the article: Mann,

M.E., Ammann, C.M., Bradley, R.S., Briffa, K.R., Crowley, T.J., Hughes, M.K., Jones, P.D., Oppenheimer, M., Osborn, T.J., Overpeck, J. T., Rutherford, S., Trenberth, K.E., Wigley, T.M.L., On Past Temperatures and Anomalous Late 20th Century Warmth, *Eos*, 84, 256–258, 2003. This figure demonstrates the lack of evidence for any periods in earlier centuries that are comparable in terms of evidence for sychronous warmth to the late 20th century. This same conclusion was also demonstrated by the recent article in *Science*: Bradley, R.S., M.K. Hughes and H.F. Diaz., Climate in Medieval Time. *Science*, 302, 404–405, 2003.

Question 8. Do you claim 22 proxies to be a sufficient sample of observations for reconstructing a Northern Hemisphere temperature? If not, why did you consider it sufficient for the 1400–1450 interval in your 1998 Nature paper? If you do, there are 29 proxies that continue to 1984 in the data base you used for your 1998 paper. Why then did you terminate your temperature reconstruction at 1980? What efforts have you made to extend the proxy re-constructions up to the present?

Response. The question is wrongly premised. The Mann et al (1998) study made use of almost 100 proxy series over the interval AD 1400–1450. The question appears to confuse the number of proxy series that was used, with the number of statistical indicators that were used to represent these proxy data. For example, the 70 series that make up the North American International Tree Ring Data Base date back to 1400, were represented in terms of their leading patterns of variance through a procedure known as "Principal Component Analysis". These patterns represented, however, a much larger number of underlying data. Most of the proxy records used in that analysis ended by 1980, limiting the useful upper limit to the calibration period used. A more recent paper (in press) extends proxy-based hemispheric temperature reconstructions through the mid-1990's, demonstrating the ability of the reconstruction to capture the accelerated warming evident in the instrumental record since 1980.

Question 9. What are the patterns of temperature change in all proxies after 1980?

Response. It is unclear what is meant by the question. Every proxy series which extends past 1980 exhibits its own particular pattern. A recent paper (in press), as referred to in Question No. 8, demonstrates that a composite of proxy temperature indicators with reliable low-frequency variability that are available through the mid-1990's capture the accelerated warming after 1980.

All of the data used in our study have been available since July 2002 on the public ftp site: ftp://holocene.evsc.virginia.edu/pub/MBH98/.

Question 10. Do you have any external (not derived by you) method or data to provide verification of your temperature reconstruction Please explain.

Response. We used the method of cross-validation to independently demonstrate the statistical reliability of our reconstructions. This is detailed in MBH98 and MBH99. We did not derive the method of cross-validation—it is a well established statistical procedure, detailed in many introductory level statistics textbooks.

Question 11. Are you aware of any errors in your data compilation for MBH98 or MBH99? If so, what are they?

Response. We are not aware of any errors. We are, however, aware of recent spurious claims of such errors by the authors of an article published in the social science journal "Energy and Environment". These claims have already been widely discredited by a cursory analysis of the paper, and a manuscript detailing the numerous fundamental errors made in the Energy and Environment paper has been submitted to the peer-reviewed literature. We would be happy to provide a copy of the paper to be made part of the official Senate record once it is published.

Question 12. Are you aware of any errors in any calculations that you made in MBH98 or MBH99? If so, what are the errors?

Response. See response to question 11.

Question 13. Vegetation grows as a result of a number of factors, including energy input, moisture supply, fire frequencies, and species competition. Do you claim it is possible to accurately remove the effects of these factors from your tree ring proxy datasets to produce a resulting time-series represents fluctuations in only air temperature? What is the magnitude of the error introduced in developing a procedure to remove these other effects? Please detail the analyses and list peer-reviewed works that specifically outline techniques to remove the effect of these other indicators for inferring past temperatures.

Response. One of the co-authors of MBH98 (Malcolm Hughes) is among the world's foremost experts in dendroclimatology, so the team of MBH98 hardly needs to be informed of the processes that influence tree growth. The method of MBH98

does not “remove” various factors from tree ring proxy information (which would be a most unwise approach!) but, rather, uses multivariate statistical methods similar to those commonly used in climate and paleoclimate field reconstruction [see e.g. Cook, E.R., K.R. Briffa, and P.D. Jones, Spatial Regression Methods in Dendroclimatology: A Review and Comparison of Two Techniques, *International Journal of Climatology, 14*, 379–402, 1994; Smith, T.M., R.W. Reynolds, R.E. Livezey, and D.C. Stokes, Reconstruction of Historical Sea Surface Temperatures Using Empirical Orthogonal Functions, *Journal of Climate*, 9, 1403–1420, 1996; Kaplan, A., Y. Kushnir, M.A. Cane, and M.B. Blumenthal, Reduced space optimal analysis for historical data sets: 136 years of Atlantic sea surface temperatures, *Journal of Geophysical Research*, 102, 27835–27860, 1997] to separate the information in the data that can meaningfully be related to surface temperature variations from that related to other influences.

Question 14. Define the difference between variability and error in a statistical analysis. In EOF analyses, is the variation of the first principal component indicative of the uncertainty associated with the data? Why or why not?

Response. Variability in an estimated quantity can be thought of as representing both ‘signal’ (the physical quantity one is interested in and ‘noise’ (everything else). The definition of noise and signal depends on a number of assumptions regarding the nature of the process that generated the times series of interest and the specification of the statistical model for the data in question. Uncertainty, which is associated with the partitioning of data variance into “noise and signal”, as defined above, depends on such detailed considerations. There are no general statistical principles that I am familiar with that relate uncertainty, thusly defined, to the first, or any other, principal component of a dataset containing both signal and noise contributions. Uncertainty is typically diagnosed by the analysis of residual variance from a statistical model based on a combined calibration/cross-validation procedure. Introductory text books such as “Statistical Methods in the Atmospheric Sciences” (D. Wilks, Academic Press) deal with this topic in detail.

Question 15. Specifically, how do you construct regional patterns of temperature changes in past centuries when data are limited, either spatially, temporally, or both?

Response. Our methods are described in detail in the following peer-reviewed scientific publications, which I would like to have made part of the official Senate record:

- Mann, M.E., Jones, P.D., Global Surface Temperatures over the Past two Millennia, *Geophysical Research Letters*, 30 (15), 1820, 10.1029/2003GL017814, 2003.
- D’Arrigo, R.D., Cook, E.R., Mann, M.E., Jacoby, G.C., Tree-ring reconstructions of temperature and sea level pressure variability associated with the warm-season Arctic Oscillation since AD 1650, *Geophysical Research Letters*, 30 (11), 1549, doi: 10.1029/2003GL017250, 2003.
- Mann, M.E., Rutherford, S., Bradley, R.S., Hughes, M.K., Keimig, F.T., Optimal Surface Temperature Reconstructions Using Terrestrial Borehole Data, *Journal of Geophysical Research*, 108 (D7), 4203, doi: 10.1029/2002JD002532, 2003.
- Rutherford, S., Mann, M.E., Delworth, T.L., Stouffer, R., Climate Field Reconstruction Under Stationary and Nonstationary Forcing, *Journal of Climate*, 16, 462–479, 2003.
- Mann, M.E., Large-scale climate variability and connections with the Middle East in past centuries, *Climatic Change*, 55, 287–314, 2002.
- Cook, E.R., D’Arrigo, R.D., Mann, M.E., A Well-Verified, Multi-Proxy Reconstruction of the Winter North Atlantic Oscillation Since AD 1400, *Journal of Climate*, 15, 1754–1764, 2002.
- Mann, M.E., Rutherford, S., Climate Reconstruction Using ‘Pseudoproxies’, *Geophysical Research Letters*, 29 (10), 1501, doi: 10.1029/2001GL014554, 2002.
- Mann, M.E., Large-scale Temperature Patterns in Past Centuries: Implications for North American Climate Change, *Human and Ecological Risk Assessment*, 7 1247–1254, 2001.
- Mann, M.E., Climate During the Past Millennium, Weather (invited contribution), 56, 91–101, 2001.
- Cullen, H., D’Arrigo, R., Cook, E., Mann, M.E., Multiproxy-based reconstructions of the North Atlantic Oscillation over the past three centuries, Paleoceanography, 15, 27–39, 2001.
- Mann, M.E., Gille, E., Bradley, R.S., Hughes, M.K., Overpeck, J.T., Keimig, F.T., Gross, W., Global Temperature Patterns in Past Centuries: An interactive presentation, *Earth Interactions*, 4–4, 1–29, 2000.
- Delworth, T.L., Mann, M.E., Observed and Simulated Multidecadal Variability in the Northern Hemisphere, *Climate Dynamics* 16, 661–676, 2000.

• Mann, M.E., Bradley, R.S. and Hughes, M.K., Northern Hemisphere Temperatures During the Past Millennium: Inferences, Uncertainties, and Limitations, *Geophysical Research Letters*, 26, 759–762, 1999.
• Mann, M.E., Bradley, R.S., and Hughes, M.K., Global-Scale Temperature Patterns and Climate Forcing Over the Past Six Centuries, *Nature*, 392, 779–787, 1998.

Question 16. Do you claim that the instrumental temperature record is known without error? If not, what error and uncertainty would you associate with the annual Northern Hemisphere averaged air temperature for 1900? For 1950? For 2000? How were these estimates incorporated into your analysis?

Response. The claim made by Dr. Legates in his testimony that we present the instrumental record without uncertainty is incorrect. If Legates, for example, were familiar with studies of the instrumental surface temperature record, he would understand that the uncertainties in this record during the 20th century are small compared to the uncertainties shown for our reconstruction [see e.g. Figure 2.1b in Folland, C.K., Karl, T.R., Christy, J.R., Clarke, R. A., Gruza, G.V., Jouzel, J., Mann, M.E., Oerlemans, J., Salinger, M.J., Wang, S.-W., Observed Climate Variability and Change, in *Climate Change 2001*: *The Scientific Basis*, Houghton, J.T., *et al.* (eds.), Cambridge Univ. Press, Cambridge, 99–181, 2001.]. Furthermore, all scientists with a proper training in statistics know that uncertainties add "in quadrature". In other words, you have to square them before adding them. This means that the relatively small uncertainty in the instrumental record makes a relatively small contribution to the total uncertainty. Legates claimed in his testimony that including the uncertainty in the instrumental record, which he estimates as 0.1oC would change the conclusions expressed by us and other mainstream climate scientists that the 1990's are the warmest decade in at least the past 1000 years within estimated uncertainties. This claim is very misleading for several reasons, First, the standard error in Northern Hemisphere mean annual temperatures during the 1990's is far smaller than the amount cited by Legates [see again Folland et al, 2001 cited above]. Even more problematic, however, Legates claim indicates a fundamental misunderstanding of the statistical concepts of standard error and uncertainty. The shaded region shown along with the Mann et al reconstruction (and other similar plots shown in recent articles such as the aforementioned "*Eos*" article, and the IPCC report) indicates two standard error intervals. The decade of the 1990's is roughly two standard errors warmer (i.e., about 0.4oC)] than any decade prior to the 20th century in the reconstruction. Based on a one-sided test for anomalous warmth, this translates to a roughly 97.5 percent level of significance. Modifying the uncertainties to include the small additional contribution due to uncertainties in the instrumental record itself would modify this only slightly, and would not lower the significance level below the 95 percent level. Though there is no such thing as an absolute estimate of uncertainty, despite Legate's implications to the contrary, a 95 percent confidence is often adopted as an appropriate criterion for significance. Legates' statement that including instrumental contributions to the uncertainty would change the conclusions is thus clearly false.

Question 17. Assuming a proxy record extended back to 1000 A.D., what specifically would be required to disqualify this proxy record from your analyses? Provide supporting evidence where others have disqualified such records from temperature analyses on these criteria.

Response. It is unclear what type of "analyses" are being referred to here. I have used a variety of different statistical methods and data in various published studies describing paleoclimate reconstruction, so the question as worded is implicitly vague. In any case, our approaches do not "disqualify" proxy data. They use objective statistical criteria to evaluate the strength of the signal available for reconstruction of the particular climate field or index (be it related to surface temperature, atmospheric circulation, drought, or other variables) to be reconstructed. Such statistical approaches, and the related approaches used by other climate researchers, are described in the various publications listed above in the response to Questions No. 13 and 15.

Question 18. Have you made available via FTP the coefficients developed to relate proxies to principal components? If not, would you make those coefficients available at NGDC/paleo?

Response. The question is based on two false premises. The first involves a naive view of what is required and expected of scientific researchers. It is unprecedented in my experience for any scientist to post in the public domain every single computational aspect of a complicated analysis. The methods of our study were adequately described in our paper and supplementary information, and the data used were made available in the public domain. Indeed, we made far more of our results,

data, and methodological details available in the public domain than is provided in most similar scientific studies. The scientific funding agencies (DOE, NSF, and NOAA) would have informed us if we had not followed the appropriate protocols in the provision of data and results.

The second false premise is a technical one. A proper understanding of the methodology employed by MBH98 would reveal that there is no one fixed set of "coefficients" that relate a particular proxy record to a particular principal components. The relationship is determined based on time-dependent inverse problem for which the weights on different records are not fixed over time, as described in our published articles.

Question 19. Have you made available via FTP any specialized computer studies, such as Matlab scripts, in connection with your temperature reconstruction? If not, would you make any such scripts used in developing the temperature reconstructions in MBH98 and MBH99 available through NGDC/paleo?

Response. The methodologies have been described, and other climate researchers have independently, successfully implemented the methodology, e.g.: Zorita, E., F. Gonzalez-Rouco, and S. Legutke, Testing the Mann *et al.* (1998) Approach to Paleoclimate Reconstructions in the Context in a 1000-Yr Control Simulation with the ECHO-G Coupled Climate Model, *Journal of Climate*, 16, 1378–1390, 2003.

Question 20. Do you claim your method of reconstructing past temperature from proxies is the only correct one? If not, please submit some published papers that use methods you consider to be correct as well. If you do consider yours the only correct method, can you provide a list of names of scientists whom you have contacted to tell them they are using the wrong methods in their work?

Response. The question is based on the false premise that my colleagues and I use any one particular "method" of reconstructing past temperatures from proxy data. In fact, I have published on the application of at least five fundamentally independent methods for using proxy data to reconstruct past climate patterns in the peer-reviewed literature. Examples of the applications of different methods can be found in the following peer-reviewed scientific publications:

- Zhang, Z., Mann, M.E., Cook, E.R., Alternative Methods of Proxy-Based Climate Field Reconstruction: Application to the Reconstruction of Summer Drought Over the Conterminous United States back to 1700 From Drought-Sensitive Tree Ring Data, *Holocene*, in press, 2003.
- Mann, M.E., Jones, P.D., Global Surface Temperatures over the Past two Millennia, *Geophysical Research Letters*, 30 (15), 1820, 10.1029/2003GL017814, 2003.
- Mann, M.E., Rutherford, S., Bradley, R.S., Hughes, M.K., Keimig, F.T., Optimal Surface Temperature Reconstructions Using Terrestrial Borehole Data, *Journal of Geophysical Research*, 108 (D7), 4203, doi: 10.1029/2002JD002532, 2003.
- D'Arrigo, R.D., Cook, E.R., Mann, M.E., Jacoby, G.C., Tree-ring reconstructions of temperature and sea level pressure variability associated with the warm-season Arctic Oscillation since AD 1650, *Geophysical Research Letters*, 30 (11), 1549, doi: 10.1029/2003GL017250, 2003.
- Mann, M.E., Bradley, R.S., and Hughes, M.K., Global-Scale Temperature Patterns and Climate Forcing Over the Past Six Centuries, *Nature*, 392, 779–787, 1998

On occasion, there are approaches used that are not adequate. For example, the approach of simply counting papers and not properly defining what constitutes an anomaly, as was the case for the paper by Soon *et al.* (2003), is most decidedly not adequate. Also, the analysis approach used by McIntyre and McItrick (2003) in which the authors attempted to reproduce the results of the previous study of MBH98 based on an analysis which used neither the same data (the authors eliminated the majority of data used by MBH98 for the first two centuries of the reconstruction) , or method as the original authors, was woefully inadequate. In fact, this latter study was described as "seriously flawed" and "silly" in a recent article in USA Today ("Global Warming Debate Heats Up Capitol Hill", 11/19/03). When deeply flawed studies such as this are published, I am interested in determining what errors have been made and, if necessary as in this latter case, promptly submitting a rebuttal to the peer-reviewed scientific literature to ensure that the scientific community is not misled by the use of inadequate approaches. To my knowledge, I am not considered to be shy in offering criticism where criticism is due.

Question 21. If there are other acceptable methods, did you try any of them on your data set prior to its publication to see what the results would be? If so would you please submit the results. If not, have you done so since? Why do you claim your multi-proxy results represent a "robust consensus," as you said in your *Eos* publication, if you have not verified that its results would also be obtained using other acceptable methods?

Response. As demonstrated both in the *Eos* article, and the various references provided in my response to Question 4, about a dozen different recent estimates based on a variety of data and approaches, published by different groups, yield statistically indistinguishable histories of Northern Hemisphere mean temperature changes in past centuries. I define such a result as characterizing a “consensus”.

Question 22. Did you at any time prior to publication compute the analysis up to 1984 or later? What were the results? If you did not, even though you had sufficient data, why not? If you did but you did not use those results, explain why. If the results were different, where did you publish a discussion of those differences? If they were the same, why did you delete them? Why, in other words, did you throw out data for the period of maximum interest?

Response. Most of the proxy records used in MBH98 and MBH99 ended by 1980, limiting the useful upper limit to the calibration period used. A more recent paper (in press) extends proxy-based hemispheric temperature reconstructions through the mid-1990’s, demonstrating the ability of the reconstruction to capture the accelerated warming evident in the instrumental record since 1980. We would be happy to provide a copy of this paper to be made part of the official Senate record when it is formally published.

Question 23. On your web site *http://www.ngdc.noaa.gov/paleo/ei/data—supp.html* where you explain the assembling of the data base for your 1980 paper you say: “Small gaps have been interpolated. If records terminates lightly before the end of the 1902–1980 training interval, they are extended by persistence to 1980.” Does this mean you made up some observations to fill in blank spots in the data records? Have you ever provided a complete public listing of all the data you made up? Please provide such a listing now. Of the 112 proxies, in how many of them did you fill gaps? Why in some of them but not others? What is the longest interval of time over which you filled in missing observations?

Response. Extension of missing values by ‘persistence’ of the final available value is a typical statistical approach to estimating small amounts of unavailable data at the end of a time series (see e.g. the textbook by Wilks, referred to in the response to Question No. 14). The fact that this approach was used to infill a modest number of missing observations between 1972 and 1980 was described in the *Nature* supplementary information. All of the data used in our study have been available since July 2002 on the public ftp site: ftp://holocene.evsc.virginia.edu/pub/MBH98/.

Question 24. What was the effect on your results of filling in the missing data? Did you run your analysis without it? Please submit the results when the filled-in data are dropped from the analysis. If it changes your results, where is that discussed If it makes no difference, why did you do it?

Response. The use of infilled data has essentially no effect on the reconstruction, as demonstrated by the fact that the same result is achieved if a 1902–1971 calibration period (which predates the use of any infilled proxy data) is used instead of a 1902–1980 calibration period. It is advisable to use the full 1902–1980 calibration interval, however, because the increased statistical constraint provided by the lengthening of the calibration period more than offsets the impact of the use of a modest amount of in-filled data in a small number of series.

Question 25. Do you agree that statistical methods based on linear extrapolation from data representing the far extreme of the line are associated with an added error/uncertainty? If so, how was this incorporated into the assessment of the error/uncertainty in your temperature reconstructions? Please provide citations from your publications. If not, please explain why the uncertainty envelope of a linear regression grows larger as a function of the distance from the mean of the data used to fit the parameters and why this was not included in your research.

Response. The so-called “leverage effect” which the question appears to refer to, is taken into account through consideration of the spectrum of the calibration residuals, allowing for resolution of any enhancement of uncertainty as a function of frequency (see MBH99). Alternatively, the uncertainties can be evaluated from an independent sample (i.e., cross-validation, rather than calibration, residuals) that eliminates any influence of calibration period leverage in the estimation of uncertainties. Both approaches give similar results [e.g. Rutherford, S., Mann, M.E., Osborn, T.J., Bradley, R.S., Briffa, K.R., Hughes, M.K., Jones, P.D., Proxy-based Northern Hemisphere Surface Temperature Reconstructions: Sensitivity to Methodology, Predictor Network, Target Season and Target Domain, *Journal of Climate*, submitted, 2003].

Question 26. Please describe the peer review process that took place with respect to your Forum article that appeared in *EOS* on July 8, 2003. If, according to the

AGU, the *EOS* Forum contains articles stating a personal point of view on a topic related to geophysical research or the relationship of the geophysical sciences to society, how can you claim that your article is peer reviewed?

Response. The article was independently reviewed and evaluated for suitability for publication by an editor who has expertise in the particular subject area. The associated process is correctly described as "peer review". Appropriate to the relatively short and non-technical nature of *Eos* "Forum" pieces, the associated peer review process is not as extensive as that employed for articles in the more technical literature such as *Geophysical Research Letters*, or *Journal of Geophysical Research*. I would suggest that the questioner contact representatives at AGU for more details on the peer-review process employed for their different journals and paper categories.

Question 27. Do you claim that producing estimates of past climate states is an exact science If so, explain why different authors can get such significantly different results when investigating and reconstructing past temperature, and detail the errors that other authors must have made. If not, explain how there can be, as you put it in your *EOS* article, a "robust consensus" regarding the correct estimate of the climate state of the past millennium.

Response. The term "exact science" is generally not used, or considered meaningful or appropriate by scientists, as science almost always involves the testing of hypotheses based on the use of intrinsically uncertain data or observations. Consistent with this fundamental aspect of nearly all scientific endeavors, my colleagues and I, and other researchers in the paleoclimate community, typically interpret the results of paleoclimate reconstructions within the context of sometimes substantial associated uncertainties. When a large number of estimates agree with each other within estimated uncertainties, and those uncertainties are modest enough to still allow for non-trivial conclusions (for example, that late 20th century warmth is anomalous in a long-term context), those conclusions can be considered as both "robust" and a "consensus".

Question 28. Please describe the peer review process that took place with respect to your 1999 *Geophysical Research Letters* paper. What were the criticisms or improvements suggested by the referees? Why was no reference made to the anomalous global warming caused by the very strong El Niño event of 1997–98 in your paper? Is this 1999 paper continuation of your 1998 paper in Nature where you stopped your reconstruction at AD 1400?

Response. The comments of reviewers on a manuscript are considered a confidential matter, involving the editor, reviewers, and authors. Providing these comments for public record would be ethically questionable, and probably violates the confidentiality policies of the associated journals. Minor suggestions were made by the reviewers and editor, and addressed to their satisfaction prior to the acceptance and publication of the paper.

Question 29. In Mann and Jones 2003 *Geophysical Research Letters*, did you change your methodology in the reconstruction of the hemispheric or global scale temperature from your prior publications? If so, why did you, and what is the rationale for the change of approach?

Response. The question is wrongly premised, as it presumes, through the use of the language "change your methodology" that scientists only have one particular methodological approach that can be applied to a problem at hand. As discussed in my answer to question 20, my research has involved the use of a variety of different methods for reconstructing past climate patterns from proxy data. The paper by Mann and Jones (2003), for example, uses a coarser resolution proxy dataset than MBH98/MBH99 and a compositing methodology that allows for the reconstruction of decadal, but not annual, changes, and the reconstruction of hemispheric mean, but not spatially resolved, patterns of temperature in past centuries. In doing so, the study was able to make use of a more restricted set of temperature records available over a longer timeframe than those used in previous high-resolution proxy reconstructions of hemispheric temperature change.

Question 30. Did IPCC carry out any independent programs to verify the calculations that you made in MBH98 or MBH99? If so, please provide copies of the reports resulting from such studies.

Response. It is distinctly against the mission of the IPCC to "carry out independent programs", so the premise of the question is false. However, the IPCC's author team did engage in a lively interchanges about the quality and overall consistency of all of the papers as the chapter was drafted and revised in the course of review.

Question 31. Did IPCC carry out any independent quality control on the data that you used in MBH98 and MBH99? If so, please provide copies of the reports resulting from such studies.

Response. The IPCC doesn't "carry out studies", so the premise of the question is false. The IPCC instead depends that the normal scientific peer-review process, especially when done in a leading journal, has ensured an acceptable level of quality. In addition, the IPCC does check to see if any criticisms have been raised postreview in comments and response to the journal articles.

Question 32. Did IPCC carry out any studies to validate the statistical procedures and methodologies used in MBH98 and MBH99? If so, please provide copies of the reports resulting from such studies.

Response. The IPCC doesn't "carry out studies", so the premise of the question is false. Instead, as indicated above, the IPCC relies on earlier stages of review to cover such matters.

Question 33. Has any organization other than IPCC or your associates carried out any independent programs to verify the calculations that you made in MBH98 or MBH99? If so, please provide copies of the reports resulting from such studies.

Response. I know of no "organizations" that carry out "independent programs" to verify calculations of individual co-authors. If the question is, have other scientists reproduced the basic results of MBH98 and MBH99, the answer is yes. Numerous other groups (see the dozen or so independent estimates of various groups shown in Figure 1 of: Mann, M.E., Ammann, C.M., Bradley, R.S., Briffa, K.R., Crowley, T.J., Hughes, M.K., Jones, P.D., Oppenheimer, M., Osborn, T.J., Overpeck, J.T., Rutherford, S., Trenberth, K.E., Wigley, T.M.L., On Past Temperatures and Anomalous Late 20th Century Warmth, *Eos*, 84, 256–258, 2003) have produced reconstructions that are remarkably similar to those of MBH98 based on a variety of data and methods. Refer back to my answer to question 4 for further details. I would like to see each of these papers made an official part of the Senate record.

Question 34. Has any organization other than IPCC conducted independent quality control on the data that you used in MBH98 and MBH99? If so, please provide copies of the reports resulting from such studies.

Response. The IPCC doesn't "carry out studies", so the premise of the question is false. The data used by MBH98 (and MBH99) were produced by other researchers, not Mann and colleagues. It is thus not clear what kind of "independent quality control" is being referred to here. However, it is fair to say that each of these papers has been subject to rigorous peer review in a leading scientific journal, which is considered by scientists to be an independent quality control process. We are aware of no criticisms of the datasets in the peer-reviewed scientific literature.

Question 35. Has any organization other than IPCC carried out any studies to validate the statistical procedures and methodologies used in MBH98 and MBH99? If so, please provide copies of the reports resulting from such studies.

Response. The IPCC doesn't "carry out studies", so the premise of the question is false. If the question were asked: Have other independent groups tested the methodology of Mann et al (1998) in a publication in the peer-reviewed climate literature, the answer would be "yes". I would refer the questioner to the following paper: Zorita, E., F. Gonzalez-Rouco, and S. Legutke, Testing the Mann *et al.* (1998) Approach to Paleoclimate Reconstructions in the Context in a 1000-Yr Control Simulation with the ECHO-G Coupled Climate Model, *Journal of Climate*, 16, 1378–1390, 2003.

The paper arrives at the conclusion that the methodology of MBH98 performs well with networks of data comparable to those used by MBH98.

Question 36. Have you ever received any communications that suggested that there might be computational errors in MBH98 or MBH99? Please provide such communications together with any responses.

Response. I receive many emails, often from list-serves of self-professed "climate skeptics" making numerous spurious claims against my work and that of many of my colleagues. I have received no correspondence providing credible evidence of any errors in our work. Nor has any such credible evidence been published in the peer-reviewed scientific literature.

Question 37. Did the peer reviewers for Nature in MBH98 carry out any independent quality control or validation studies? If so, please provide copies of such reports.

Response. Neither I, nor authors of peer-reviewed journal articles in general, are made privy to the detailed analyses that peer reviewers may or may not have performed in the process of reviewing a manuscript. Authors only receive the comments

that were selected to be made available to them by the reviewer and editor. This question is thus impossible to answer. Numerous other groups (see the dozen or so independent estimates of various groups shown in Figure 1 of: Mann, M.E., Ammann, C.M., Bradley, R.S., Briffa, K.R., Crowley, T.J., Hughes, M.K., Jones, P.D., Oppenheimer, M., Osborn, T.J., Overpeck, J.T., Rutherford, S., Trenberth, K.E., Wigley, T.M.L., On Past Temperatures and Anomalous Late 20th Century Warmth, *Eos*, 84, 256–258, 2003) have produced reconstructions that are remarkably similar to those of MBH98 based on a variety of data and methods. See my answer to Question No. 4.

Question 38. 38. Did the peer reviewers for *Geophysical Research Letters* in MBH99 carry out any independent quality control or validation studies? If so, please provide copies of such reports.

Response. See response to Question No. 37.

Question 39. How many people have requested the underlying digital information in MBH98? Please provide dates of such requests and dates of your reply.

Response. My collaborators and I have not kept a specific record. The data has been provided to any scientific groups that have requested it, and has been made available on an open access basis through a public ftp site: ftp://holocene.evsc.virginia.edu/pub/MBH98/, since July 2002.

Question 40. Were you one of the primary or lead authors of IPCC/TAR chapter 2?

Response. The convening lead authors of chapter 2 of the IPCC TAR were Dr. Chris Folland and Thomas Karl. I was one of eight additional co-authors contributing to chapter 2.

Question 41. In your capacity as IPCC/TAR author, did you prepare any drafts that referred to your own papers? Please provide all drafts that you prepared for IPCC.

Response. I contributed to numerous sections of the chapter and provided contributions that referenced the work of the leading paleoclimatologists, which includes me and many of my colleagues. Those interested in drafts of IPCC chapters should inquire of the appropriate IPCC working group. I am not in possession of such drafts, and even if I were, I would not be at liberty to distribute the various drafts of the chapters of the report.

Question 42. Was any language from your drafts referring to your own reports ultimately used by IPCC/TAR? Please provide highlighted versions from IPCC.

Response. The wording of the question is unclear. If the question is, did I, in my contributions to the chapter, provide summaries that included references to my own work as well as that of other scientists, the answer is of course yes. Since each of the authors was asked to contribute sections related to their particular areas of expertise, and since the IPCC authors were chosen from among the leading scientists in the world, it would be distinctly odd if it were not the case that most authors referred to their work, as well as that of others, in their contributions.

Question 43. Did IPCC/TAR have any policies governing how lead authors used their own work? Did IPCC/TAR have any quality control procedures in the event that a lead author used his own work? Please provide a short summary of your understanding of such procedures.

Response. I am not a spokesperson for the IPCC. However, it is my understanding that the IPCC carries out a process for developing its summarization of the understanding of science that leads to one of the most rigorously peer-reviewed scientific documents in existence. Individual technical chapters are prepared by expert scientific teams that consider the full range of published papers in a subject area. This expert author team then solicits an initial peer review from a large number of other scientists in the field, drawing on those representing the full range of expert science. The reports next go through a much wider review that is open to literally thousands of scientists around the world. Finally countries, NGO's, and professional groups (such as business groups) are provided the opportunity to send in review comments. (and in the case of the U.S. government review, an invitation to submit comments to be considered to be forwarded to the IPCC is published in the Federal Register, enabling all to participate in this review). With the comments available at each stage of the review process, the authors consider each comment and document their response. The meticulousness and fairness of the revision process by the authors in response to reviewer comments is evaluated by an independent pair of "review editors" who are themselves top international climate scientists who are not authors of the report itself. The National Academy of Sciences, at President George W. Bush's request, and other national academies around the world have independently

reviewed the process and the validity of the scientific findings of the IPCC and endorsed them.

Question 44. Did MBH98 and MBH99 use any proxy series, which were either unpublished or which resulted from unpublished calculations, which you carried out? If so, please identify, and detail how you verified those unpublished results.

Response. MBH98 and MBH99, as many studies, made use of newly available data that had not yet been published by the original authors providing those data, and thus was provided to Mann and colleagues on a provisional basis that they not release the data until the authors had a chance to publish the records themselves. After all of the data used had been published, the full dataset used by MBH98 and MBH99 was made available in the public domain on the public website: ftp://holocene.evsc.virginia.edu/pub/MBH98/

Question 45. Despite solar variability over the last two millennia, your analysis concludes the Northern Hemisphere average temperature has remained virtually constant. What mechanism or mechanisms are responsible for negating the influence of the sun? Do climate models (GCMs) exhibit the same lack of response to solar forcing that your analysis implies? If not, why are model simulations at variance with your conclusions and how does that limit their applicability for future climate scenario assessments?

Response. The question is falsely premised on several levels. No reasonable description of the reconstructions that we or others have produced of temperature variations in past centuries would characterize them as “virtually constant”. The reconstructions performed by my group and others indicate an amplitude of variability that consistent with expectations from models driven with estimates of past radiative forcing including solar and radiative forcing, and allowing for the added role of internal unforced variability [see e.g. Crowley, T.J., Causes of Climate Change Over the Past 1000 Years, *Science*, 289, 270–277, 2000]. Indeed, it has been shown that the model-predicted pattern of surface temperature response to solar forcing in past centuries closely resembles that estimated from the temperature reconstructions that my colleagues and I have performed [Shindell, D.T., Schmidt, G.A., Mann, M.E., Rind, D., Waple, A., Solar forcing of regional climate change during the Maunder Minimum, *Science*, 294, 2149–2152, 2001; Waple, A., Mann, M.E., Bradley, R.S., Long-term Patterns of Solar Irradiance Forcing in Model Experiments and Proxy-based Surface Temperature Reconstructions, *Climate Dynamics*, 18, 563–578, 2002; Shindell, D.T., Schmidt, G.A., Miller, R., Mann, M.E., Volcanic and Solar forcing of Climate Change During the Pre-Industrial era, *Journal of Climate*, in press, 2003].

Question 46. How did the temperatures of the mid-Holocene Optimum Period (6000 to 9000 BP) compare with those observed today? Was it a global or a local phenomenon? What was or were the cause or causes of any temperature anomalies in that period? What is the cause of the 104 to 105 year timescale changes in deuterium, oxygen isotope, etc., concentrations in ice core records? Are such changes global or local?

Response. Paleoclimate experts have established that mid-Holocene warmth centered roughly 5000 years ago was restricted to high latitudes and certain seasons (summer in the Northern Hemisphere and winter in the southern hemisphere). Because much of the early paleoclimate evidence that was available (for example, fossil pollen assemblages) came from the Northern Hemisphere extratropics, and is largely reflective of summer conditions, decades ago some scientists believed that this was a time of globally warmer conditions. It is now well known that this is not the case. More abundant evidence now demonstrates, for example, that the tropical regions were cooler over much of the year. All of these changes are consistent with the expected response of surface temperatures to the known changes in the Earth’s orbital geometry relative to the Sun during that time period and associated climate feedbacks, as detailed in peer-reviewed scientific publications [e.g., Hewitt, C.D., A Fully Coupled GCM Simulation of the Climate of the Mid-Holocene, *Geophysical Research Letters*, 25 (3), 361–364, 1998; Ganopolski, A., C. Kubatzki, M. Claussen, V. Brovkin, and V. Petoukhov, The Influence of Vegetation-Atmosphere-Ocean Interaction on Climate During the Mid-Holocene, *Science, 280*, 1916–1919, 1998].

Climate model simulations indicate quite good agreement with paleoclimate evidence now available. These models calculate that global annual average temperatures were probably about the same or a few tenths of a degree C cooler than today (the late 20th century) during this time period [Ganopolski, A., C. Kubatzki, M. Claussen, V. Brovkin, and V. Petoukhov, The Influence of Vegetation-Atmosphere-Ocean Interaction on Climate During the Mid-Holocene, *Science*, 280, 1916–1919, 1998; Kitoh, A., and S. Murakami, Tropical Pacific Climate at the mid-Holocene and the Last Glacial Maximum simulated by a coupled oceanatmosphere general circula-

tion model, *Paleooceanography, 17* (3), (19)1–13, 2002.]. That's a far cry from the very out-of-date claim made by Dr. Legates in his testimony. Dr. Legates' comments regarding climate changes over the past 1000 years reflect a similar lack of familiarity with a whole body of paleoclimate research, especially with the new insights gained through the augmented research program, during the past decade.

Question 47. It has been observed that in the past, carbon dioxide concentrations have sometimes lagged air temperature trends; that is, changes in air temperature have subsequently sometimes resulted in changes in carbon dioxide concentrations. Do you agree with those results from expert researchers? Why or why not?

Response. The question mis-characterizes the evidence that has been provided by paleoclimate researchers. The studies that the questioner appears to be alluding to, demonstrate a phase relationship between ice core CO_2 estimates and *local* temperature variations at the site of the ice core. Furthermore these local temperature estimates are indirectly inferred from oxygen isotopes, based on quite uncertain assumptions regarding oxygen isotope paleothermometary and neglecting possible biases due to the variable seasonality of local accumulation. As local temperature variations at the site of the ice core have an unknown relationship with global mean temperature variations (which are far more dominated by lower latitudes which occupy the majority of the Earth's surface area), the phase relationships between past CO_2 and global mean temperature variations are not known. In spite of these qualifications, it is not at all implausible that the geologic record indicates that at some times the CO_2 increase may lag the initial temperature increase; such a situation would be expected, for example, if the change in climate was initiated by a change in the orbital geometry that affected the distribution of solar radiation, and then the slow warming drove CO_2 from the warming ocean into the atmosphere. It is because of the many possibilities for how different processes can interact that it is essential to not simply base a conclusion on an apparent correlation without evaluating the underlying physical mechanisms for that particular period.

Question 48. Are there any time periods for which atmospheric CO_2 content has changed without a concomitant change in global air temperature? Are there periods when the atmospheric CO_2 content was relatively high butglobal air temperatures relatively low?

Response. In his testimony, Dr. Legates indicated that there were historical cases where the temperature has gone up, but that CO_2 has fallen. It may well be the case that this has happened in the past. However, it is hardly surprising, and certainly not inconsistent with our established understanding of the various factors that influence surface temperatures. The warming response to increased greenhouse gas concentrations lags the actual increase in greenhouse gas concentrations in the atmosphere potentially by several decades, due to the sluggish response of the oceans, which have an enormous thermal capacity compared to the atmosphere, to increased surface radiative forcing. So warming is not expected to be contemporaneous with changes in CO_2, but instead, to lag it by several decades. However, greenhouse gases are certainly not the only factor affecting the average surface temperature of the Earth. There are other anthropogenic factors, such as increased sulphate aerosols, which can have a cooling effect on the climate, and natural factors, such as volcanic activity, modest natural variations in solar output, and internal dynamics associated with climate events such as El Niño, which also influence the average surface temperature of the globe. At any particular time, these other factors may outweigh the warming effect due to increased greenhouse gases. For example, the relative lack of warming during the period 1940–1970 appears to be related to a combination of such factors, as discussed in my response to an earlier question. But while these other factors tend to cancel over time, the increased greenhouse gases lead to a systematic warming that will not cancel out over a very long time period. It is for precisely this reason that late 20th century warming now appears to have risen above the range of the natural variability of past centuries.

Question 49. Two independent and nearly direct measures of surface temperature (deep borehole reconstructions) over the past several millennia have been published for Greenland (Dahl-Jensen et al 1998) and the Middle Urals (Demeshko and Shchapov 2001). The local surface temperature at these locations is highly correlated with global temperature on 10-year time scales and longer (r2 > 50 percent 10 yr with agreement increasing for longer averaging periods). Both reconstructions independently show their local surface temperatures were at least a1 °C warmer for century-scale mean temperatures around A.D. 900 than the latter half of the 20th century, translating into a global anomaly of at least +0.2° C relative to today.

This further implies that even higher global temperature anomalies for shorter periods, such as half-century or decadal periods, were observed about 1000 years

ago. Why do these two robust measures of local and global approximations differ greatly from Mann *et al.* 1999?

Response. The question is wrongly premised on multiple levels. First, the correlations cited are completely wrong. No citation to the peer-reviewed literature is provided, so it is difficult to determinate how these numbers were arrived at. I therefore proceeded to analyzed the appropriate surface air temperature gridpoint data from the Climatic Research Unit of the University of East Anglia myself. I found that only after 1922 is there adequate coverage (>50 percent areal coverage) to estimate a meaningful Greenland areal-mean temperature. For the period back to 1922, the linear correlation between the Greenland and Northern Hemisphere mean temperature is r= -0.06 (negative!), nor is there a significant correlation at decadal or longer timescales. In fact, the trends in the two series during the latter 20th century are of opposite sign. So the numbers cited are completely spurious.

It is in fact well known by the climate community that there are fundamental physical reasons *why* temperatures in Greenland are, in general, poorly correlated with Northern Hemisphere mean temperature. Owing to the strong overprint of processes, such as the North Atlantic Oscillation, and changes in coupled ocean-atmosphere processes in the North Atlantic that impart a large regional overprint of temperature variation in this region, both negative and positive correlations with Northern Hemisphere mean temperature can be found, depending on the time period and region of Greenland analyzed.

The Dahl Jensen et al Greenland borehole data may indeed be useful temperature proxy data for the regions they represent, and they have been used in reconstructions of Northern Hemisphere mean temperature, with caveats due to their extremely low temporal resolution (see Mann and Jones, 2003). While the two Greenland borehole records show significantly different histories over the past 1000 years (which is expected since temperature trends vary markedly depending on the region of Greenland in question), one of the two records does correlate well with the instrumental Greenland record over the period of mutual overlap. Its shows the mid-20th century warm peak, followed by the latter 20th century cooling peak, just as the instrumental Greenland annual mean temperature record does. However, instrumental Northern Hemisphere mean temperature has, in contrast, *warmed* markedly during the latter 20th century.

The Greenland borehole temperature reconstruction may tell us something about temperatures in Greenland over the past few millennia even though the two different Greenland borehole records show some differences between them. But these results are unlikely to tell us much, if anything, about Northern Hemisphere mean temperature trends. Indeed, Dahl-Jensen et al have never, to my knowledge, claimed in their studies that that temperature variations in the two regions of Greenland reconstructed (which themselves show significantly different histories over the past 1000 years) are representative of Northern Hemisphere mean temperatures, and I would be surprised if the authors were comfortable in having their data represented as such.

Question 50. In your view what should the Federal Government do in response of rising concentrations of CO_2? What would be the climate impact of this effort?

Response. In my view, the Congress and the Federal Government should be taking the scientific findings of the mainstream research community very seriously and should stop focusing so much attention on the poorly conducted and distracting nitpicking of the various contrarian scientists. The IPCC assessments represent the most authoritative reviews of the science and have been unanimously endorsed by all of the participating nations of the world—it is time to pay attention to their findings. Exactly what steps should be taken and how fast this should be done are policy questions that members of this body should be responsibly and thoughtfully addressing. The long residence timescales of anthropogenic greenhouses gases, and the lags in the response of the climate system (e.g. sea level rise) to already realized increases in greenhouse gas concentrations dictate, however, that there are potentially significant costs to delayed action.

Question 51. Approximately what percentage of the temperature increase in the observational record over the last 100 years would you attribute to anthropogenic causes? What percentage would you attribute to increased urbanization? What percentage would you attribute to non-urbanized land use changes? What percentage would you attribute to natural (solar, volcanic, etc.) variability? What percentage would you attribute to "internal" climate variability? What percentage would you say results from other or unexplained sources? Give estimates for the years 1900,1940, 1980 and 2000.

Response. A cursory review of the available evidence (see e.g. Figure 2.1 of chapter 2 of the 2001 IPCC Scientific Working Group report) indicates the following ap-

proximate attributes in the observed record of global mean temperature changes over the past 100 years: a warming of approximately 0.3° C to 1940, a statistically insignificant change (given the uncertainties) from 1940 to the mid-1970's, and then an additional warming of approximately 0.5° C from 1970 to 2000. This pattern of behavior is reproduced closely by models driven with estimates of both natural and anthropogenic forcing of the climate during the 20th century. The period of relative stasis in global mean temperatures from 1940 to 1970, in these model simulations, appears to result from the cooling impact of anthropogenic aerosols (for which there was a large increase during that time period) as well as a potential cooling contribution from explosive volcanic eruptions that occurred during that period, which tended to offset the warming influence of increased greenhouse gas concentrations during that time period (e.g. the 1957 eruption). However, much of the overall warming of the globe during the 20th century (which is between 0.6° C and 1.0° C depending on the precise instrumental data set used, and the precise endpoints of the interval examined) is clearly a result of increased greenhouse gas concentrations, as established in these simulations, consistent with the conclusions of the IPCC Third Assessment Report that most of the warming of the past 50 years is attributable to human influences.

Question 52. What was the earth's climate like the last time the atmospheric concentration of carbon dioxide was near today's level of about 370 parts per million (ppm) and what were past conditions like when concentrations were at 550 ppm? Detail the factors that cause the global carbon cycle to produce these high levels of atmospheric carbon dioxide.

Response. It is not precisely known what the "earth's climate" was like the last time carbon dioxide levels were near 370 ppm (let alone 550 ppm) because the available paleoclimate evidence available this long ago are quite uncertain and incomplete. That having been said, it is believed, based on the available proxy information and faunal/floral evidence, that global temperatures were probably several degrees higher than they are today when CO_2 concentrations neared 550 pm, roughly consistent with model simulation results. One probably has to go back roughly 40–50 million years ago (see chapter 3 of the 2001 IPCC working group 1 report) to find a time when CO_2 concentrations were in the range of 550 ppm (i.e., roughly double their pre-industrial concentration) and approximately 80 million years ago (i.e., the mid-Cretaceous period when Dinosaurs roamed the polar regions) to find a time when CO_2 levels were in excess of 1200 ppm (a level that will be reached, at current rates of CO_2 increase, within 1 to 1½ centuries). Proxy evidence available for this period, tenuous though it is, suggests deep ocean temperatures 8–12° C warmer than present. State of the art climate model simulations performed by Bette Otto-Bleisner and colleagues using the National Center for Atmospheric Research (NCAR) global climate model, which incorporate such CO_2 levels (and the continental configuration corresponding to the mid-Cretaceous period), indicate significantly warmer sea surface temperatures, with tropical sea surface temperatures approximately 4° C warmer and polar sea surface temperatures approximately 6–14° C warmer than present. The simulations indicate an absence of perennial sea ice at even the most polar latitudes.

Question 53. In your vitae, you indicate that you serve on the panel for NOAA's Climate Change Data and Detection (CCDD) program, while at the same time, you also have received large grants from this program. Please explain your role on the panel, how grant submissions are evaluated, and why there is no conflict of interest or impropriety associated with members of a panel receiving large grants from the program for which they serve.

Response. I am not a spokesperson for NOAA, and would suggest that the questioner contact the appropriate NOAA agency officials for further information on their conflict of interest and disclosure policies. That notwithstanding, however, I would note the following points. Government funding agencies seek to draw upon the leading experts of the field in their panels. Inevitably, this means that specific science programs within NSF and NOAA invite to their review panels scientists who typically submit proposals themselves to those panels. Scientists are also asked to disclose any conflicts of interest they might have in reviewing a proposal, and are asked to recuse themselves from any participation in discussions related to proposals that they might have even peripheral involvement with. In my involvement in both NSF and NOAA panel reviews, I have on many occasions recused myself from reviewing or discussing a proposal based on such considerations.

Question 54. Do you receive any income or reimbursement (travel, speaking fees, etc.) from any sources, which have taken advocacy positions with respect to the Kyoto Protocol, the U.N. Framework Convention on Climate Change, or legislation

before the U.S. Congress that would affect greenhouse gas emissions? If so, please identify those sources and the approximate amount of money that they represent.

Response. All income or travel expense reimbursement funds that I have received to my recollection have come from academic institutions, government funding agencies such as NSF, NOAA, NASA, DOE, and scientific organizations such as the American Geophysical Union and University Corporation for Atmospheric Research (UCAR). I am not familiar with any advocacy positions that have been taken by any of these institutions or organizations regarding the U.N. Framework Convention on Climate change, or legislation before the U.S. Congress that would affect greenhouse gas emissions.

RESPONSES BY MICHAEL MANN TO ADDITIONAL QUESTIONS FROM SENATOR JEFFORDS

Question 1. Is it your understanding that during the mid-Holocene optimum period (the period from 4000–7000 B.C.) that annual mean global temperatures were more than a degree C warmer than the present day?

Response. This is an oft-repeated but patently false claim. Dr. Legates, who has no established expertise in the relevant field of paleoclimatology, indeed asserts that temperatures were warmer at this time. In fact, not only is that not the consensus of the paleoclimate research community, but just the opposite is believed to be true of global annual mean temperatures at this time. Paleoclimate experts know that the mid-Holocene warmth centered roughly 5000 years ago was restricted to high latitudes and certain seasons (summer in the Northern Hemisphere and winter in the southern hemisphere). Because much of the early paleoclimate evidence that was available (for example, fossil pollen assemblages) came from the Northern Hemisphere extratropics, and is largely reflective of summer conditions, decades ago some scientists believed that this was a time of globally warmer conditions. It is now well known that this is not the case. More abundant evidence now demonstrates, for example, that the tropical regions were cooler over much of the year. All of these changes are consistent with the expected response of surface temperatures to the known changes in the Earth's orbital geometry relative to the Sun during that time period and associated climate feedbacks, as detailed in peer-reviewed scientific publications [e.g., Hewitt, C.D., A Fully Coupled GCM Simulation of the Climate of the Mid-Holocene, *Geophysical Research Letters, 25* (3), 361–364, 1998; Ganopolski, A., C. Kubatzki, M. Claussen, V. Brovkin, and V. Petoukhov, The Influence of Vegetation-Atmosphere-Ocean Interaction on Climate During the Mid-Holocene, *Science, 280*, 1916–1919, 1998].

Climate model simulations indicate quite good agreement with paleoclimate evidence now available. These models calculate that global annual average temperatures were probably a few tenths of a degree C cooler than today during this time period [Kitoh, A., and S. Murakami, Tropical Pacific Climate at the mid-Holocene and the Last Glacial Maximum simulated by a coupled ocean-atmosphere general circulation model, *Paleooceanography, 17* (3), (19)1–13, 2002.]. That's a far cry from the very out-of-date claim made by Legates. Legates' comments regarding climate changes over the past 1000 years reflect a similar lack of familiarity with a whole body of paleoclimate research, especially with the new insights gained through the augmented research program, during the past decade.

Question 2. Why only focus on the past 1000 or 2000 years and not further back?

Response. Large changes in climate certainly occurred in the distant past. If we look million years back in time, dinosaurs were roaming the polar regions of the Earth, and the globe was several degrees wanner than today. Carbon dioxide levels were probably several times their current level, slowly having attained such high levels due to changes in the arrangements of the continents ('plate tectonics') which influence the volcanic outgassing of carbon dioxide from the solid Earth. These changes occurred on timescales of tens of millions of years. Going back 10,000 years ago, large ice sheets existed over North America due to natural changes that occur in the Earth's orbit around the Sun on timescales of tens of thousands of years. Trying to study distant past climates for insights into modern natural climate variability is hampered by the fact that the basic external constraints on the system (the continental arrangement, the geometry of the Earth's astronomical orbit, the presence of continental ice sheets—what we call the 'boundary conditions') were significantly, different from today. Focusing on the evolution of climate in the centuries leading up to the 20th century (i.e., the past 1000 to 2000 years) provides a perspective on the natural variability of the climate prior to the period during which large-scale human influence is likely to have occurred, yet modern enough that the basic boundary conditions on the climate system were otherwise the same. This provides us, in essence, 'control' for diagnosing whether or not recent changes in climate are

indeed unusual. Moreover, only during the past 1000–2000 years do we have adequate networks of proxy climate data with the required (annual) resolution in time to compare and validate against modern instrumental records; Reliable quantitative reconstructions of large-scale surface temperature patterns further back in time are thus not, at present, possible.

Question 3. One of the Northern Hemisphere temperature reconstructions in your Figure 1 (the green curve from a paper by *Esper* and colleagues from *Science* in 2001) shows larger swings in past centuries, marginally outside the uncertainty bounds of the other reconstructions and model simulations. Does this indicate internal inconsistency in our knowledge?

Response. There is no inconsistency. *Esper et al* noted, in their paper, that their estimate, unlike that of my colleagues and mine (the Mann/Bradley/Hughes or "MBH" reconstruction), was not representative of the entire Northern Hemisphere. They explicitly noted this in their paper, where they pointed out the likely reason for differences is that the MBH reconstruction represents the full Northern Hemisphere (tropics, subtropics, and extratropics) while the *Esper et al* reconstruction only represents the restricted extratropical continents. In fact, in a *Science* article that appeared in the same issue as the *Esper et al* paper [Briffa, K.R. and T.J. Osborn, *Science*, Blowing Hot and Cold, 295, 2227–2228, 2002] Briffa and Osborn noted that much of the difference was due to an arguably inappropriate scaling that *Esper et al* used, and an inappropriate comparison of summer vs. annual temperatures; "when we regressed the record of Esper *et al.* against non-smoothed data (see the figure), this difference (with MBHJ was reduced to about 0.4° C. Recalibrating both curves against year-by-year warm season temperatures reduces this difference further to about 0.35° C."

As shown in the "*Eos*"article discussed in my testimony, which represents a consensus of the leading researchers in the field [Mann, ME, Ammann, C.M., Bradley, R.S., Briffa; K.R., Crowley, T.J., Hughes, M.K.; Jones, P.D., Oppenheimer, M., Osbom, T.J., Overpeck, J.T., Rutherford, S., Trenberth, K.E., Wigley, T.M.L., On Past Temperatures and Anomalous Late 20th Century Warmth, *Eos*, 84, 256–258, 2003], a proper scaling of *Esper et al* record prior to comparison with other estimates, shows it only marginally outside the error estimates of the MBH reconstruction and the many other estimates that are in agreement with it. As noted in two articles in *Science* that shortly followed the *Esper et al* paper [Mann, M.E., Hughes, M.K., Tree-Ring Chronologies and Climate Variability, *Science*, 296, 848, 2002; Mann, ME, The Value of Multiple Proxies, *Science*, 297, 1481–1482, 2002], it is likely that the emphasis of the *Esper et al* reconstruction on only the summer season and the extratropical continental regions, provides a biased estimate of the true pattern of annual, hemisphere-wide temperatures in past centuries, explaining the small differences between this estimate and other estimates. This conclusion has been verified in recent modeling studies [Shindell, D.T., Schmidt, G.A., Miller, R., Mann, M.E., Volcanic and Solar forcing of Climate Change During the Pre-Industrial era, *Journal of Climate*, in press, 2003]. It is thus clearly disingenuous when contrarians make the argument that the *Esper et al* result is in conflict with the mainstream conclusions of the climate research community with regard to the history of Northern Hemisphere mean annual temperature variations over the past millennium as embodied, for example, in the *Eos* article.

Question 4. As climatologist, can you explain what kind of quantitative analysis it takes to determine whether or not the last 50 years has been unusually warm compared to the last 1000 years?

Response. Well, such an analysis requires the careful use of proxy data, because we don't have widespread instrumental temperature records prior to the mid-19th century. By 'careful use': I mean that one must first establish that the records actually resolve the changes of the past 50 years. This typically requires annually resolved proxy records or the very circumspect use of records with decadal resolution. One must not, as in the Soon and Baliunas' studies, use records that do not resolve the trends of the past few decades. One must also establish the existence of an actual temperature signal in the available proxy data before using them to reconstruct past temperature patterns, and one must properly synthesize regional data, which typically all show different tendencies at any given time, into an estimate of the average temperature over the entire hemisphere or globe. There area number of ways of performing such a synthesis, from the sophisticated pattern reconstruction approaches my colleagues and I have described in the technical literature, to the relatively straightforward compositing approach that many other paleoclimatologists, (including myself) have also used. In all cases, the estimates based on the proxy data must be calibrated against modern instrumental temperature measurements, to allow, for a quantitative estimate of past temperatures. The estimate must then

be independently verified or, what we call, 'cross-validated,' by showing that it independently reproduces earlier instrumental data that were not used to calibrate the estimates. Finally, uncertainties must be diagnosed based on how well the reconstruction describes actual available instrumental measurements. Once such steps have been taken, it is possible to compare the recent instrumental record to the reconstruction within the context of the uncertainties of the reconstruction. This latter comparison allows us to gauge whether or not late 20th century temperatures are anomalous or not in a long-term context. The conclusion from legitimate such studies that late 20th century warmth is indeed anomalous in a millennial or longer-term context has been shown to be quite robust with respect to the details of the data set used, or the methodology used (as shown in exhibit 1 in my testimony, the first figure of the "*Eos*" piece). It is noteworthy that the Soon and Baliunas paper satisfies none of the required standards for a 'careful use of proxy data' specified above.

Question 5. Do you claim that appropriate statistical methods do not exist for calibrating statistical predictors, including climate proxy records, against a target variable, such as the modern instrumental temperature record?

Response. No. The statement belies a centuries-old field of statistics known as "multivariate linear regression" in which a set of candidate "predictors" (such as proxy data) are statistically related to a target variable or "predictand" (such as the instrumental temperature record) during a common interval of overlap (e.g., the 20th century). If done properly, this statistical method isolates the temperature information that is contained within the proxy data, and uses that information to reconstruct past temperature patterns from the proxy data. It is also well known to those properly trained in statistics that the "regression model" must be independently "validated" by showing that it successfully reproduces independent data (e.g. longer-term instrumental temperature records) that were not used in constructing the statistical model itself. The estimates by Mann and colleagues embrace each of these fundamental statistical principles.

Legates in his testimony seemed to claim that a composite estimate (e.g. of Northern Hemisphere mean temperature) should somehow resemble *each* of the individual predictors (e.g. the various regional temperature estimates). Such a result is, in general, a statistical impossibility. There is a basic theorem in statistics known as the "central limit theorem" which indicates a general tendency for a composite (i.e. average) of a large number of different individual estimates to cancel out in terms of the pattern variation and amplitude of variability evident in the individual estimates. If Legates were indeed (as he claims to be) familiar with the instrumental surface temperature record, he would know, for example, that individual instrumental thermometer records available for particular locations over the globe during the 20th century show very little in common with the 'composite' series constructed by averaging all of the individual records into a hemispheric or global estimates. Because season-to-season and year-to year fluctuations in the climate at regional scales often result from shifts in the atmospheric circulation, not every location experiences the same variation in any given year; for example, this summer, the western United States and Europe are anomalously warm, while the East Coast is anomalously cool. The average series reflects a tendency for a cancellation of the various "ups" and "downs" in the different individual series that often occur at different times. This is simply a statement about the instrumental record itself, and requires no use of proxy data at all. It is discouraging that Legates and colleagues haven't performed even such simple analyses with available instrumental data that would expose the fundamental flaws in their supposed 'statistical' reasoning.

In this context, it should be noted that Legates testimony seriously misrepresents the statistical analyses used by Mann and co-workers. In his testimony, Legates claimed that my collaborators and I replaced the proxy data for the 1900's by the instrumental. The assertion is simply factually incorrect. Any reader of our published work knows full well that our proxy-based temperature reconstruction extends well into the late 20th century, through 1980 (the vast majority of published high-resolution climate proxy data are not available at later times than this). It is shown in our work that the reconstruction independently reproduces estimates from the instrumental surface temperature data record back through the mid-19th century and, in certain regions, back through the mid-18th century (i.e., our regression model is "validated" in the manner discussed above). Our Northern Hemisphere average temperature series reconstructed from proxy data is shown to agree with the instrumental Northern Hemisphere average record over the entire interval available for comparison (1856–1980). This successful validation of the reconstruction, furthermore, allows us to compare the proxy-based reconstruction to the entire instru-

mental record (available through 1999 at the time of our publication), taking into account the uncertainties in the reconstruction.

Legates further claim that we present the instrumental record without uncertainty is disingenuous. If Legates, for example, were familiar with studies of the instrumental surface temperature record, he would understand that the uncertainties in this record using the 20th century are minimal compared to the uncertainties shown for our reconstruction [see e.g. Figure 2.1b in Folland, C.K., Karl, T.R., Christy, J.R., Clarke, R. A.,Gruza, G.V., Jouzel, J., Mann, M.E., Oerlemans, J., Salinger, M.J.;, Wang, S.-W., Observed Climate Variability arid: Change, in *Climate Change 2001: The Scientific Basis*, Houghton, J.T., *et al.* (eds.), Cambridge Univ. Press, Cambridge, 99–181, 2001.]. Furthermore, all scientists with a proper training in statistics know that uncertainties add "in quadrature", In other words, you have to square them before adding them. This means that the relatively small uncertainty in the instrumental record makes a relatively small contribution to the total uncertainty. Legates claimed in his testimony that including the uncertainty in the instrumental record, which he estimates as 0.1° C would change the conclusions expressed by us and other mainstream climate scientists that the 1990's are the warmest decade in at least the past 1000 years within estimated uncertainties. This claim is very misleading for several reasons, First, the standard error in Northern Hemisphere mean annual temperatures during the 1990's is far smaller than the amount cited, by Legates [see again Folland et al, 2001 cited above]. Even more problematic, however, Legates claim indicates a fundamental misunderstanding of the statistical concepts of standard error and uncertainty,. The shaded region shown along with the Mann et al reconstruction (and other similar plots shown in recent articles such as the aforementioned "*Eos*" article, and the IPCC report) indicate two standard error intervals. The decade of the 1990's is roughly two standard errors warmer (i.e., about 0.4° C) than any decade prior to the 20th century in the reconstruction. Based on a one-sided test for anomalous warmth, this translates to a roughly 97.5 percent level of significance. Modifying the uncertainties to include the small additional contribution due to uncertainties in the instrumental record itself would modify this only slightly, and would not lower the significance level below the 95 percent level. Though there is no such thing as an absolute estimate of uncertainty, despite Legate's implications to the contrary, a 95 percent confidence is often adopted as an appropriate criterion for significance. Legates statement that including instrumental contributions to the uncertainty would change the conclusions is thus clearly false.

Question 6. In determining whether the temperature of the "Medieval Warm Period" was warmer than the 20th century, does your work analyze whether a 50-year period is either warmer or wetter or drier than the 20th century? Is it appropriate to use indicators of drought and precipitation directly to draw inferences of past temperatures?

Response. No, the work of me and my colleagues does not follow the flawed approach used by Soon and Baliunas. It is fundamentally unsound to infer past temperature changes directly from records of drought or precipitation. The analysis methods used by my various collaborators and I (e.g. the 13 authors of the recent article in *Eos*) employ standard statistical methods for identifying the *surface temperature* signal contained in proxy climate records, and using only that temperature signal in reconstructing past temperature patterns. By contrast, Soon and Baliunas simply infer evidence for warm "Medieval Warm Period" or cold "Little Ice Age" conditions from the relative changes in proxy records often reflective of changes in precipitation or drought, rather than temperature. It is difficult to imagine amore basic mistake than misinterpreting hydrological evidence in terms of temperature evidence. This fundamental shortcoming in their approach is identified in the *Eos* article referred to in my testimony that was written by 13 leading climate and paleoclimate scientists. Incidentally, the *Eos* piece was peer-reviewed, despite the claim by Legates otherwise in his testimony—an associate editor with training in the particular field corresponding to the submitted piece ("Atmospheric Science" in this case); reviews the content of *Eos* forum pieces prior to acceptance.

Question 7. Can you compare the quantitative analysis that supports your conclusion that the climate is warming faster in recent years than at any time in the recent past with the analysis done in the Soon literature review?

Response. Technically speaking, there is no actual "analysis" in the Soon and Baliunas review, in that they don't appear to have performed a single numerical or statistical operation upon a single time series at all. They do not provide any quantitative estimates of temperature changes, let alone any estimate of uncertainties, Instead, they claim to interpret the results of past studies mainly by counting the number of studies coming to some conclusion; no matter how they got there and

whether there have been later interpretations. Science requires analysis—not just counting studies. Climate scientists whose records they analyzed have gone on record as indicating that Soon and Baliunas misinterpreted their studies (e.g., the article by David Appell in the August 2003 issue of *Scientific American*) and numerous climate scientists have indicated (see same article) that Soon and Baliunas misinterpreted evidence of drought or precipitation as evidence of temperature changes, did not use records that resolve the climate changes of the late 20th century, and did not take into account whether or not variations in different regions were coincident or not. Soon and Baliunas also neglect undeniable evidence of substantial warming of hemispheric and global surface temperatures during the past few decades. So the Soon and Baliunas analysis fails just about every meaningful criterion that might be applied to determining the validity of an analysis that purports to evaluate current warming in the context of past temperature trends. This deeply flawed study thus contrasts sharply with other rigorous quantitative studies (as discussed in the *Eos* article) performed by numerous other scientists with appropriate training in the fields of climatology and paleoclimatology, which use proper statistical methods for inferring past temperature changes from proxy data, provide uncertainty estimates, and employ appropriate comparisons of current and past trends.

Question 8. What was the earth's climate like the last time that atmospheric concentrations of carbon dioxide were at today's levels or about 370 parts per million (ppm) and what were conditions like when concentrations were at 550 ppm, which will occur around 2060 or so?

Response. We have to go back far into the past to find carbon dioxide levels approaching today's levels. Ice core studies indicate that modern carbon dioxide levels are unprecedented for at least four glacial/inter-glacial cycles: in other words, for more than 400,000 years. Other evidence suggests that carbon dioxide levels are now higher than they have been for at least 10 million years. One probably has to go back roughly 40–50 million years ago (see chapter 3 of the 2001 IPCC working group 1 report) to find CO_2 concentrations were in the range of 550 ppm (i.e., roughly double their preindustrial concentration) and approximately 80 million years ago (i.e., the mid-Cretaceous period when Dinosaurs roamed the polar regions) to find CO_2 levels in excess of 1200 ppm (a level that will be reached, at current rates of CO_2 increase, within one to one-and-a-half centuries). Proxy evidence available for this period, tenuous though it is, suggests deep ocean temperatures 8–12° C warmer than present. State of the art climate model simulations performed by Bette Otto-Bleisner and colleagues using the National Center for Atmospheric Research (NCAR) global climate model, which incorporate such CO_2 levels (and the continental configuration corresponding to the mid-Cretaceous period), indicate significantly warmer sea surface temperatures, with tropical sea surface temperatures approximately 4° C warmer and polar sea surface temperatures approximately 6–14° C warmer than present. The simulations indicate an absence of perennial sea ice at even the most polar latitudes.

Question 9. Are you aware of any scientists beside the authors of the Soon et al article who support using "wetness" or "dryness" as indicators of past temperatures, instead of actual temperatures or proxy data that reflects temperatures?

Response. I am not aware of any other scientist who has made the mistake of interpreting paleoclimatic information in this way. As discussed above, trained paleoclimatologists typically use statistical methods to identify the strength of the temperature signal in a proxy record prior to using it in reconstructing past temperature patterns.

Question 10. Is there any known geologic precedent for large increases of atmospheric CO_2 without simultaneous changes in other components of the carbon cycle and the climate system?

Response. There is not, to my knowledge, such an example. As discussed above, the geological record shows a clear relationship between periods of high CO_2 and relatively high global mean temperatures. The study of the relationship between changes in CO_2 and climate in the paleoclimate record is sometimes complicated by the fact that these relationships can be relatively complex during rapid transitions between glacial and interglacial climates such as those that occurred with the coming and going of ice ages that occurs on a roughly 100 thousand year timescale over the past nearly one million years. However, one can turn to periods of time when the climate and CO_2 were not varying rapidly, and thus the climate was approximately in an "equilibrium" state, for insights into the relationship between CO_2 and climate, A perfect such example is the height of the last ice age, the so-called "Last Glacial Maximum" or "LGM" centered roughly 25 thousand years ago. At this time, CO_2 was substantially lower than today (just below 200 ppm) and global mean sur-

face temperatures were several degrees (about 4C or so) cooler than today. Such relationships between past CO_2 changes and global temperature changes typically, suggest a "sensitivity" of the climate system to enhanced CO_2 of 1, 5 to 4.5 C warming for each doubling of CO_2 concentrations from their pre-industrial levels, similar to the range of sensitivities found in various climate models.

Question 11. According to a study published in Science magazine recently [B. D. Santer; M. F. Wehner, T. M. L. Wigley, R. Sausen; G. A. Meehl, K. E. Taylor, C. Ammann, J. Arblaster, W. M. Washington, J. S. Boyle, and W. Brüggemann Science 2003 July 25; 301; 479–483], manmade emissions are partly to blame for pushing outward the boundary between the lower atmosphere and the upper atmosphere. How does that fit with the long-term climate history and what are the implications?

Response. This is yet another independent piece of evidence confirming a detectable anthropogenic influence on climate during the late 20th century. This evidence is consistent with evidence for unprecedented surface warming during the past few decades—warming that indeed appears unprecedented for as long as we have records (i.e., for at least a thousand, and probably two thousand years). These changes, moreover, are consistent with predictions from climate models driven by known anthropogenic (human) forcing of the climate.

Question 12. At this hearing, there were a number of calls for "sound science." Could you please explain what it is about the IPCC process that justifies respecting the IPCC results as the very soundest representation of the science of climate change?

Response. The IPCC carries out a process for developing its summarization of the understanding of science that leads to one of the most rigorously peer-reviewed scientific documents in existence, Individual technical chapters are prepared by expert scientific teams that consider the full range of published papers in a subject area. This expert author team then solicits an initial peer review from a large number of other scientists in the field, drawing on those with the full range of views. The reports next go through a much wider review that is open to literally thousands of scientists around the world. Finally countries, NGOs, and professional groups (such as business groups) are provided the opportunity to send in review comments. At each stage, authors consider each comment and document their response. The meticulousness and fairness of the revision process by the authors in response to reviewer comments is evaluated by an independent pair of "review editors" who are themselves top international climate scientists who are not authors of the report itself. The National Academy of Sciences, at President George W. Bush's request, and other national academies around the world have independently reviewed the process and the validity of the scientific findings of the IPCC and endorsed them. To question the IPCC and the IPCC process, as does Dr. Legates thus not only does a disservice to thousands of the world's top scientists; but to the exceptional care and rigor of the process that has led to the unanimous adoption of all of the IPCC's assessments by representatives of the over 150 nations that participate in the IPCC process. The documents are very finely honed and carefully phrased. The scientific studies of those such as or, Legates are considered, as are their review comments, and it is terribly disingenuous, not to mention totally unacceptable to the international community, after all of the care and consideration put into these efforts to try to so cavalierly dismiss them.

In his testimony Legate's alleges that the IPCC report misrepresents what is known about climate change in past centuries and that it somehow replaces conventional wisdom with dramatically new conclusions. One must conclude that Legates either did not read the report, or if he did, he did not understand what he read; for if he had he would certainly have to recognize factually incorrect nature of his comments. The IPCC chapter dealing with paleoclimatic evidence discussed the full range of regional evidence described in the peer-reviewed as well as evidence from hemispheric composites that average the information from different regions. The paper by Soon and Baliunas is a dramatic throwback to the State of our knowledge many decades ago, while the IPCC report provides a far more upto-date assessment of all of the available knowledge regarding past climate change. The Soon and Baliunas papers provide a glaring example of the very "unsound" science that Senator Inhofe claims to be concerned about, as numerous mainstream climate researchers have now opined in the media, and in the peer-reviewed scientific literature.

Question 13. In your opinion, how do the processes used by the IPCC and the National Academy of Sciences compare to the process used in the publication of the Soon and Baliunas paper and other papers by so-called "contrarians"? In the next IPCC assessment, would you expect that the Soon and Baliunas paper will be con-

sidered and cited, putting it into the context of other papers and findings and explaining why it has differences or similarities?

Response. As discussed above, the IPCC is one of most rigorously peer-reviewed scientific documents in existence. By contrast, the contrarians rarely publish in the peer-reviewed literature and when they do, it is not uncommon to discover, as in the case of the Soon and Baliunas paper, irregularities in the peer-review process. Publisher Otto Kinne indeed indicated that the review process at Climate Research "failed to detect methodological flaws" in the Soon and Baliunas paper. I would indeed expect that in the next IPCC assessment, the Soon and Baliunas paper will be discussed and evaluated in the context of other available evidence and in the context of how it is faring in the literature when the IPCC review takes place (e.g., if there are as many criticisms about their work as at present, and they have not been seriously addressed by more careful followup studies by the authors. I would think their conclusions will be rejected as scientifically unsound. I would not presume to know in detail what the result of the assessment will be, but I believe it fair to assume that the rigorous review provided by the IPCC assessment process will not fail to identify the methodological flaws that appear to have slipped through the cracks in their publication in the journal "Climate Research".

Question 14. Could you explain the IPCC's lexicon for indicating relative levels of confidence and how you would suggest this relates to the information being "real" and "probable"? When IPCC says something is "very likely," just what do they mean?

Response. To avoid the type of misunderstanding that often results from when scientists seek to convey scientific results to a non-technical audience, the IPCC specifically sought to employ a lexicon in which terms such as "likely" or "probable" or "very likely" had specific statistical meanings attached to them. A fairly conservative standard was typically employed in this process. Consider the conclusion in the IPCC report that the 1990's are the warmest decade in at least the past 1000 years for the Northern hemisphere average temperature. This conclusion is based on the fact that the average warmth of the 1990's exceeds that for any reconstructed decade in the reconstructed Northern Hemisphere series. To be more specific, the 1990's warmth exceeds any past decade by two standard errors. This corresponds to a roughly 97.5 percent probability based on standard statistical assumptions. Probabilities of 90 percent–99 percent are termed "very likely" in lexicon typically adopted by the IPCC report. However, this conclusion was offered as only "likely" (corresponding to a 66 percent–90 percent level of probability) rather than the more stringent "very likely" because it was based on only a small number of independent studies at the time. Since that time, of course, several more studies have affirmed this conclusion, and one might imagine that a more stringent conclusion will be offered in the future. This example nonetheless illustrates the manner by which IPCC adopted conservative standards in their use of terms such as "likely" or "very likely". It is instructive to contrast that standard with the one taken in the Soon and Baliunas paper. The Soon and Baliunas paper does not provide a quantitative estimate of any quantity (such as average Northern Hemisphere temperature), or any assessment of uncertainty. It is thus not possible for the authors to attach any meaningful statement of likelihood or probability to any of their conclusions. They thus provide no basis for judging the validity of any of the claims made in their paper, in striking contrast to the rigorous standards adopted by IPCC, and by the work of my collaborators and me.

Question 15. In his questioning, Senator Inhofe cited results regarding the potential costs of implementing the Kyoto Protocol from the Wharton Econometric Forecasting Association (WEFA). I realize that you are not an economist, but would you please comment as a scientist on the following two points:

(a) Senator Inhofe cited economic projections (e.g., 14 percent increase in medical costs; real income drop of $2,700 per household) going out a decade or so into the future and without any indication of uncertainty on these estimates implying an accuracy of two-significant figures. Senator Voinovich cited numbers to similar claimed accuracy going out 20–25 years; again without any indication of uncertainty. Could you please comment on what you think are the relative strengths and weaknesses of making climate projections based on use of physical laws versus economic extrapolations and what sorts of relative uncertainty should likely be associated with each type of estimate so that they can be interpreted in a comparative way by decisionmakers?

Response. Indeed, it is somewhat remarkable that politicians who reject the validity, for example, of climate model simulations, which are based on solution of the laws of physics, can so uncritically accept precise economic projections based on economic forecasts based on untestable and unverifiable assumptions governing human

decisionmaking, and speculative future scenarios that depend on the unfolding of the political process. Specific numbers of "14 percent" or "$2,700" are of course entirely dependent on the assumptions that go into such forecasts. Because assumptions about future economic growth and about future government policies are necessarily uncertain, estimates of changes in costs or income from such models must also be quite uncertain, and so should have large uncertainties (or ranges) associated with them. The faith expressed in such poorly constrained economics estimates by the same individuals who express strong skepticism of results from far more physically based and testable climate simulation models, strikes me as a remarkable inconsistency.

(b) Because you may be familiar with the 1997 study by the World Resources Institute entitled "The costs of climate protection: A guide for the perplexed," which explains the important role of assumptions in leading to very different cost estimates from even one economic model, much less among different models, could you explain in a comparative fashion how robust the findings regarding the "hockey stick" behavior of the climate in the various studies carried out: by you and fellow investigators may be to variations in the assumptions made. It's clear that varying the assumptions among reasonable possibilities in the economic models can change what is calculated to be a few percent impact on the economy to a small gain; would changes in the assumptions you are making change the indication of strong anthropogenic warming to an indication of human-induced cooling?

Response. The millennial temperature reconstruction (or "hockey stick" as it was termed by former GFDL head Dr. Jerry Mahlman) is based on a rigorously validated statistical model with demonstrated predictive skill based on comparisons with independent data. The primary conclusions drawn from the reconstruction (e.g. the anomalous nature of late 20th century warmth) are based on a conservative appraisal of the uncertainties in the reconstruction, and are not strongly dependent on the assumptions made. The same—conclusions have been affirmed now by several other independent empirical and-model-based estimates. As discussed above, this contrasts with economic predictions which are necessarily far more sensitive to the assumptions that go into them.

Question 16. In his opening statement, Senator Inhofe concluded that the Soon and Baliunas paper is "credible; well-documented; and scientifically defensible." By contrast, your testimony indicated that the experts in the field do not consider this to be the case. Does one have to be an expert in the field to understand the apparent problems with this paper? If not, could you summarize in terms for non-scientists what the key problems with the paper are?

Response. The mainstream scientific research community has indeed rejected the approach, interpretation, and conclusions advanced by Soon and Baliunas as fundamentally unsound. The major flaws in the analysis, as described in earlier comments in more detail, are basic enough that they can be understood by a non-specialist. In short, the analysis by Soon and Baliunas is unsound because (a) they inappropriately interpreted indicators of past precipitation as evidence of past temperature changes, (b) they did not use an approach which takes into account the simultaneity and lack of simultaneity of variations in different regions, and (c) they did not employ a proper standard for evaluating recent changes (i.e., changes during the past few decades) in the context of past variations. Indeed, they also misinterpreted past published work, and did not provide any quantitative estimates, let alone estimates of uncertainty. It is difficult to find anything of scientific merit at all in their published work.

Question 17. As I understand it, the data that both you and Soon and Baliunas draw from is the same, and it is not a question of the data being the problem. Instead, it is apparently the processing of the data that you are indicating has been done in a substandard way by Soon and Baliunas. Is this correct?

Response. Keep in mind that Soon and Baliunas, unlike my collaborators and I, don't actually analyze any data at all. They simply claim to have 'interpreted' past studies (often incorrectly). Soon and Baliunas refer to a number of proxy data studies that describe data which we employed in our analysis. There are also many proxy data that we used in our study that the Soon and Baliunas do not discuss. However, as alluded to in the question, the real issue isn't what data were used. Numerous independent now-published studies employing widely different assemblages of proxy climate data sets have demonstrated (e.g, as in the *Eos* article discussed earlier) a similar pattern of past variations in hemispheric mean temperature. The real issue with Soon and Baliunas is indeed not the set of studies that they claim to have interpreted, but rather the approach that they took to interpreting those studies. As indicated in previous comments, Soon and Baliunas, unlike mainstream climate researchers, did not employ a method for isolating the actual

temperature signal in the proxy records before using the records to draw conclusions regarding past temperature changes. Unlike mainstream researchers, they did not aggregate information in a way that addresses whether or not variations in different regions are simultaneous. Unlike mainstream researchers, they did not analyze the actual modern (late 20th century) warmth in the context of past variations. And finally, they did not even produce a quantitative estimate of past temperature variations, let alone an estimate of uncertainty.

Question 18. The Soon et al literature review has been described as shifting the paradigm away from the "hockey stick" description of global warming. It seems that that review simply attempted to revive an older theory of climate change that has been discarded by NOAA, the USGCRP, the NAS, the IPCC, etc. Please comment.

Response. The mainstream scientific community has clearly and decisively rejected the Soon and Baliunas papers as scientifically unsound which undermines any claims made by industry-funded special interest groups or politicians, that the papers provide any valid scientific conclusions, let alone the basis for a shift of the "paradigm". Indeed, the Soon and Baliunas papers simply promote a number of long-discredited myths which have been replaced, in recent decades by far more rigorous and quantitative analyses as described by the IPCC, USGCRP, NAS, and other mainstream scientific organizations and funding agencies. In short, the Soon & Baliunas papers simply repackage myths that were discredited more than a decade ago.

Question 19. In an opinion-editorial by former Secretary of Energy James Schlesinger, he suggested that "we have only a limited grasp of the overall forces at work, . . ." in terms of global climate change. Could you please summarize for us what the scientific community considers the key forces at work over the past 1000 years, how well these estimates are understood, and whether there is a general consistency or inconsistency between the various forcings and the climate estimates that you and colleagues have developed for the last 1000 years?

Response. Mr. Schlesinger's assertions entirely mischaracterize the nature of our scientific knowledge, which has advanced tremendously during the past several decades. In fact, a very large number of peer-reviewed scientific studies have been published in the leading scientific, journals such as *Nature* and *Science* in the past two decades elucidating the role of natural and anthropogenic factors in observed climate changes. Physically based models have been developed and validated against observations, and these models reproduce complex climate phenomena such as El Niño. These same models have been driven with the primary "external" factors that are believed to govern climate variations on timescales of decades and centuries. These external factors include natural factors, such as the modest estimated variations in radiative output of the Sun, which varies by a fraction of a percent over time, variations in the frequency and intensity of explosive volcanic eruptions, which have a several-year cooling effect on the climate through the injection of reflective volcanic aerosols into the stratosphere, and very small changes in the Earth's orbit relative to the Sun that occur on multi-century timescales. These external factors also include the "anthropogenic" influences of increased greenhouse gas concentrations due to fossil fuel burning, changes in the reflective properties of the land surface due to human land use alterations, and the regional cooling effect of anthropogenic sulphate aerosols in certain industrial regions. When driven with these factors, these climate models have demonstrated a striking ability to reproduce observed global and hemispheric temperature trends during the 20th century, as well as longer-term trends in past centuries as reconstructed from proxy data. Such results are nicely summarized in the 2001 IPCC scientific working group report. Mr. Schlesinger would have benefited from a reading of this report, or the follow-up National Academy of Sciences report that endorsed the key IPCC conclusions, prior to writing his oped piece, which reflects a decades-old understanding of the State of the science.

Question 20. In Dr. Soon's testimony, he speaks about there being "warming" and "cooling" for different periods. If he did not construct an integral across the hemisphere or a real timeline; can he say anything other than that there were some warm periods and cool periods?

Response. Aside from not adequately distinguishing temperature changes from hydrological changes, Dr. Soon and his collaborators indeed did not even attempt to estimate contemporaneous patterns of past temperature change, let alone an integral across the hemispheric domain to assess hemispheric mean temperature changes. It is unclear what, if any, meaningful conclusions can be drawn from the Soon and Baliunas study.

Question 21. Dr. Soon indicates that "local and regional, rather than global average changes are the most relevant and practical measure of climate change and its impact." Could you please comment on this, including the relative likelihood of identifying a signal of climate change amidst the local fluctuations? In what sense might local changes be the most practical measure? In that the primary forcings of the climate are global in scale, does it not make most sense to first determine how the large-scale rather than the regional climate might be affected?

Response. Dr. Soon's comments are truly misguided. Firstly, the surface temperature reconstructions published by my colleagues and I explicitly resolve regional patterns of surface temperature, so it is entirely unclear why Dr. Soon believes that we don't address regional climate changes. Had Dr. Soon understood our papers, he would be aware that we do. However, unlike the study that Dr. Soon published, our reconstructions explicitly take into account the issue of the relative timing and simultaneity of surface temperature changes in different regions. Only through doing this it possible to form an integrated measure of temperature changes such as hemispheric mean temperature. Scientists with training in climatology, statistics; and other areas of research required in the study of paleoclimate reconstruction know that the signal-to-noise ratio of any surface response to global radiative climate forcing increases as the scale of spatial averaging increases, In discussions of climate change it is thus the integral of the surface temperature field over an entire hemisphere or globe, which constitutes the most useful single variable for detecting, and attributing causal factors to observed changes. The spatial signature of the surface temperature signal (both with respect to position on the surface of the Earth, and altitude in the atmosphere) can nonetheless help to distinguish one source of climate forcing (e.g. solar) from another (e.g. enhanced greenhouse gases). My colleagues and I have indeed used the spatial patterns of surface temperature changes in past centuries to identify the role of natural external forcing of climate [Shindell, D.T., Schmidt, G.A., Mann, M.E, Rind, D., Waple, A., Solar forcing of regional climate change during the Maunder Minimum, Science, 294, 2149–2152, 2001; Waple, A., Mann, M.E., Bradley, R.S., Long-term Patterns of Solar Irradiance Forcing in Model Experiments and Proxy-based Surface Temperature Reconstructions, *Climate Dynamics*, 18, 563–578, 2002; Shindell, D.T., Schmidt, G.A., Miller, R., Mann, M.E., Volcanic and Solar forcing of Climate Change During the Pre-Industrial era, *Journal of Climate*, in press, 2003]. Both Dr. Soon and Dr. Legates advocate in their testimony a primary role of solar forcing in recent climate change, though they provide no quantitative justification for this assertion at all. In fact, nearly a dozen detailed "detection" and "attribution" studies published during the past decade by leading climate researchers in the premier international scientific peer-reviewed journals such as *Science and Nature*, have shown that the vertical and horizontal pattern of observed warming is inconsistent with the response of the climate to solar forcing, but is consistent, with the response of the climate to anthropogenic forcing. Thus a prudent use of spatial information, as described in various studies by leading climatologists, including my collaborators and I, can potentially help elucidate the roles of natural and anthropogenic factors. However, Dr. Soon's studies are deficient in their use of any such information, and provide no insights into the factors governing past climate change.

Question 22. This year, the western United States is anomalously hot and dry. The eastern United States is wetter than it has been since approximately 1891 and cool. Europe is hotter and drier than it has been in about 150 years. If we assume for the moment that these types of anomalies would persist for 50 years, are these the types of anomalies that Soon and Baliunas would consider as being indicative of there being an equivalent to the Medieval Warming in the western United States and Europe while at the same time there is the equivalent of the Little Ice Age in eastern North America? How would your type of approach vary in its analysis of the year 2003 compared to the apparently contradictory results that Soon and Baliunas would have?

Response. Indeed, as my colleagues and I discussed in our peer-reviewed articles in "*Eos*" and more recently "*Geophysical Research Letters*" [Mann, M.E., Jones, P.D., Global Surface Temperatures over the Past two Millennia, *Geophysical Research Letters*, 30 (15), 1820, doi: 10.1029/2003GL017814, 2003] the Soon and Baliunas approach is indeed internally contradictory in that it would separately identify anomalies for even a given year, such as 2003, as simultaneously supportive of conditions they would classify as associated with a "Little Ice Age" and a "Medieval Warm Period" anomaly. As outlined in the question, this year's pattern of climate anomalies is a perfect example. Trained climatologists and paleoclimatologists know that one must independently evaluate precipitation or drought information from temperature information in reconstructing past climate patterns. For example, colleagues of mine

and I have developed reconstructions of patterns of drought over the continental U.S. in past centuries from droughtsensitive tree-ring data [Cook, E.R., D.M. Meko, D.W. Stahle, and M.K. Cleaveland, Drought Reconstructions for the Continental United States, *Journal of Climate*, 12, 1145–1162, 1999; Zhang, Z., Mann, M.E., Cook, E.R., Alternative Methods of Proxy-Based Climate Field Reconstruction: Application to the Reconstruction of Summer Drought Over the Conterminous United States back to 1700 From Drought-Sensitive Tree Ring Data, *Holocene*, in press, 2003]. The drought reconstructions display a quite different pattern of behavior over time from reconstructions of Northern Hemisphere mean temperatures, just as patterns of drought over the continental U.S. during the 20th century as recorded from instrumental data show relatively little in common with instrumental Northern Hemisphere mean temperature estimates (for example, the most prominent drought episode was the 'dust bowl' of the 1930's, while the most prominent anomaly in the Northern Hemisphere temperature record is the late 20th century warming). Drought and temperature are essentially independent climate variables. The papers by Soon and Baliunas seem not to recognize this fundamental fact. Finally, there is an irony in the testimonies of Soon and Legates in that they seem to be criticizing my colleagues and me for supposedly only focusing on the reconstruction of temperature patterns, when in fact we, and not they, have published work reconstructing past patterns of drought, precipitation, and atmospheric circulation from proxy climate data. However, we have made careful use of the information contained in proxy data in independently reconstructing patterns of temperature and patterns of drought. By contrast, Soon and colleagues hopelessly convolute such information in their interpretations of past climate trends.

Question 23. Could you provide a more detailed explanation for the apparent Northern Hemisphere, cooling from the 1940's to 1970's? What is the general expectation of what would have happened to the climate in the absence of any human influences, so just continuing on from the trend for the last 1000 years prior to human intervention?

Response. In fact, this issue has been studied by quite a number of climate scientists for well over a decade. As I mentioned in my testimony, a statistically significant cooling trend from 1940's to the 1970's is not evident for the globe, but only, for the Northern Hemisphere. Dr Legates testimony on this matter is incorrect in that regard. The observed record of global-mean temperature changes over the past 100 years indicates a warming to about 1940; little change from 1940 to the mid-1970's, and then further warming, Legates implies in his comments that these changes are inconsistent with our current understanding of the factors governing climate change. This is also incorrect. In order to understand these observed changes it is necessary to consider all likely causal factors, both anthropogenic and natural, Anthropogenic factors include the warming effects of greenhouse gases and the cooling effects of sulfate aerosols. Natural factors include changes. in the output of the Sun and the effects of explosive volcanic eruptions (such as the El Chichon eruption in 1982 and the Mt. Pinatubo eruption in 1991). and internal variability associated with natural climate oscillations in the ocean circulation and various modes of coupled ocean-atmosphere variability (such as El Niño). When all of these factors are considered, models give an expected pattern of 20th century temperature changes that is in remarkable agreement with the observations—and the models clearly show the three phases noted above. In particular, the leveling off of the warming trend over 1940–1975 turns out to be explained largely by the relatively rapid increase in cooling effects of sulfate aerosols as the world emerged from the Depression and WWII (and perhaps a small contribution from natural, internal variations in ocean currents). This cooling temporarily offset the warming due to increasing concentrations of greenhouse gases. This was first pointed out in a paper by Dr. Tom Wigley of the National Center for Atmospheric Research (NCAR) in *Nature* in 1989 and has been verified by numerous additional studies since. This agreement between models and observations shows quite clearly that human factors have been the dominant cause of global-scale climate change over the past 50 years, contrary to the repeated assertions by Soon and Legates that they are a manifestation of natural climate variability. In the absence of anthropogenic factors, model simulations indicate that natural factors alone would have lead to a slight cooling trend of global temperatures over the 20th century [Crowley, T.J., Causes of Climate Change Over the Past 1000 Years, *Science, 289*, 270–277, 2000], in stark contrast to the dramatic warming that has been observed.

Question 24. It was suggested at the hearing that increased CO_2 could enhance plant life, and that since plants produce oxygen, this could lead to more O_2 and less CO_2. Could you please comment on the likelihood of this and how large the percent-

age changes could possibly be, recognizing that as the CO_2 decreased, this would presumably mean the plants would do less well and conditions would cool?

Response. Those suggestions (for example, Legates testimony with regard to the role of the 'CO_2 fertilization' effect) represent a misunderstanding of the factors governing carbon cycle dynamics and their interaction with climate. In fact, careful studies that have been performed with coupled climate/terrestrial carbon cycle models that take into account the internal coupled interactions between climate and carbon dioxide, accounting for multiple potential factors such as (a) the potential 'CO_2 fertilization' effect in which productivity of plants increases in a higher CO_2 environment, (b) the impact of climate on productivity in which higher surface temperatures favor enhanced plant growth, and (c) the feedback of CO_2 back on surface temperature alluded to in the question [see chapter 3 of the 2001 IPCC working group 1 report]. Such studies show that changes in surface temperature, through their impact on biological productivity, have led to, at most, changes of 5 to 10 ppm in CO_2—levels over the past 1000 years [see Gerber, S, Joos, F., Bruegger, P.P., Stocker, T.F., Mann, M.E.; Sitch, S., Constraining Temperature Variations over the last Millennium by Comparing Simulated and Observed Atmospheric CO_2; *Climate Dynamics*, 20, 281–299,2003]. Such changes are minimal in comparison with the dramatic increases in CO_2 concentrations of more than 80 ppm associated with human activity, suggesting that the 'CO_2 fertilization' effect advanced by Legates in his testimony in reality has a minimal role, at best, in the modern changes taking place in CO_2 concentrations and climate.

Question 25. Could you please clarify your remarks regarding the FACE experiments? When you say that increased CO_2 leads to more uptake and that they will rot, do you mean that all plants will grow and eventually die and decay, and that increased CO_2 really only ties up a bit more carbon in the process?

Response. The sequence of questions and time allotted did not allow me to adequately explain this basic, but important point. The point I was making in my testimony is that the supposed increase in the terrestrial carbon reservoir due to enhanced plant growth that is argued to occur in a higher CO_2 concentration atmosphere (the so-called 'CO_2 fertilization' effect) is not a long-term, sustained effect, It is only a short-term effect that lasts only over the generational timescales of forest stands. Any depletion of the atmospheric carbon reservoir due to enhanced growth or productivity of plants argued to arise from higher CO_2 concentrations is short lived, because the plant or tree eventually dies and gives its carbon back to the atmosphere either through microbial activity (rotting) or burning. In other words, when plants, with any potential additional organic carbon storage that might arise from enhanced biological activity eventually die, they don't simply pile up in place with their carbon reservoir intact (which is what is implicitly assumed by those who argue that 'CO_2 fertilization' represents a potential long-term offset to anthropogenic CO_2 increases).

Instead, this carbon is acted upon by biological, chemical, or physical processes which serve to add the carbon back to the atmosphere. Thus, the so-called "CO_2 fertilization" effect, cannot serve as a permanent offset to anthropogenic increases in the atmospheric carbon budget (i.e., atmospheric CO_2 concentrations), as implied by Legates in his testimony. It may simply act to slow, slightly, the rate of CO_2 increase in the atmosphere by slightly increasing the storage rate (but not the residence time) of carbon in the terrestrial biosphere.

Another way to estimate the potential influence is by considering the total amounts of carbon presently stored in vegetation. Present, about 600 billion tons of carbon are tied up in the aboveground vegetation. About 2–3 times this much is tied up in roots and below ground carbon, which is a more difficult carbon pool to augment. By comparison, scenarios for fossil fuel emissions for the 21st century range from about 600 billion tons (if we can keep total global emissions at current levels, which implies controls well beyond the Kyoto Protocol calls for) to over 2500 billion tons if the world increases its reliance on combustion of coal as economic growth and population increase dramatically. These numbers clearly indicate that sequestering a significant fraction of projected emissions in vegetation is likely to be very difficult, especially as forests are cleared to make way for agriculture and communities. While there are possibilities of storage in wells and deep in the ocean, stabilizing the atmospheric CO_2 concentration would require gathering up the equivalent of 1 to 2 times the world's existing above ground vegetation and putting it down abandoned oil wells or deep in the ocean. While CO_2 fertilization will help to increase above ground vegetation a bit, storing more than a few tens of percent of the existing carbon would be quite surprising, and this is likely to be more like a few percent of global carbon emissions projected for the 21st century.

Question 26. Senator Thomas stated that "[t}he rise in temperature during the 20th century occurred between 1900 and 1940." Could you please provide an indication of how much change occurred during this period based on internationally accepted observations, and compare this to the total change during the 20th century? Also please comment on whether it is scientifically representative to calculate a change starting with a cold period due to volcanic eruptions and end it during a period devoid of volcanic eruptions and then compare it to the century long period, which had major volcanic eruptions in both the first and last decades of the century.

Response. A cursory review of the actual evidence (see e.g. Figure 2.1 of chapter 2 of the 2001 IPCC Scientific Working Group report) indicates the following approximate attributes in the observed record of global-mean temperature changes over the past 100 years: a warming of approximately 0.3° C to 1940, a statistically insignificant change (given the uncertainties) from 1940 to the mid-1970's, and then an additional warming of approximately 0.5° C from 1970 to 2000. Senator Thomas' claim is thus clearly mistaken. As discussed in my answer to an earlier question, this pattern of behavior is reproduced closely by models driven with estimates of both natural and anthropogenic forcing of the climate during the 20th century. The period of relative stasis in global mean temperatures from 1940 to 1970, in these model simulations, appears to result from the cooling impact of anthropogenic aerosols (for which there was a large increase during that time period) as well as a cooling contribution from explosive volcanic eruptions that occurred during that period, which tended to offset the warming influence of increased greenhouse gas concentrations during that time period. However, much of the overall warming of the globe during the 20th century (which is between 0.6° C and 1.0° C depending on the precise instrumental data set used, and the precise endpoints of the interval examined) is clearly a result of increased greenhouse gas concentrations, as established in these simulations.

Question 27. Senator Thomas stated that "there is no real evidence" that the greenhouse gases are affecting the climate. Could you please summarize the available evidence explaining their probable effect? Please include in your answer a specific example of a proxy indicator such as tree rings and explain the various subtleties in deriving a temperature.

Response. As discussed in my answers to previous questions, the fact that increased greenhouse gas concentrations have a role in 20th century warming is no longer considered as being in doubt by mainstream researchers. Even noted contrarians such as Patrick Michaels of the Cato Institute now agree with this conclusion. The only room for legitimate scientific debate concerns the relative role of greenhouse gas concentrations vs. other factors, and the rate of future warming that may occur. Evidence establishing the role of anthropogenic greenhouse gas increases in 20th, century warming includes the agreement with the full spatial (horizontal and vertical) pattern of warming with predictions from model simulations, and the fact that only model simulations which include anthropogenic forcing can match the observations, as discussed earlier. Evidence for an anthropogenic influence on climate also comes from evidence of the anomalous nature of late 20th century warmth in a very long-term context (i.e., in at least the past millennium, and potentially the past several millennia or longer). One such source of evidence for this conclusion comes from proxy climate records (such as tree rings, corals, and ice cores) that can be used to reconstruct long-term temperature patterns based on a careful consideration of the temperature signal in those data, as discussed in my response to earlier questions. But other evidence of anomalous late 20th century warmth comes from indications of unprecedented melting of mountain glaciers the world over (including meltback in the Alps so dramatic that it recently revealed the now-famous "Ice Man" who had been trapped in ice for more than 5000 years), and evidence of unusual phenological changes (e.g. the timing of flowering of plants) during the late 20th century.

Question 28. Senator Carper asked the other two witnesses if they thought it "possible to emit unlimited amounts of CO_2 into our atmosphere without having any impact on climate or temperature?" What is your expectation of what would occur? That is, how much change in the CO_2 concentration would cause how much of a response?

Response. The response of global mean surface temperature to increased CO_2 varies roughly as the logarithm of the CO_2 concentration (meaning that increments in temperature scale with the percentage change in CO_2 rather than the change in amount itself). This is a very well known, and long established result that follows both from basic theoretical considerations of radiative transfer theory, and is embodied in experiments using global climate models with varying levels of CO_2 concentrations. The statistical relationship between estimated concentrations of CO_2

and the admittedly crude estimates of global mean temperatures at various periods in the geological past or during past glacial intervals, conform relatively well to this theoretical relationship within estimated uncertainties [see e.g. the textbook, "Earth's Climate Past and Future", by W.F. Ruddiman (WR Freeman and Co), 2001]. I was extremely surprised when Dr. Soon indicated that he did not know how to answer Senator Carper's question, suggesting that he is not familiar with this fairly basic scientific knowledge.

This result implies, in the absence of any other factors, a linear increase in temperature over time in response to an exponential increase in CO_2 (which is not a bad description of the character of the CO_2 trend associated with exponentially increasing anthropogenic activity over the past two centuries). Climate models tell us that the "slope" of that linear increase is between 1.5° C to 4.5° C for each doubling of the CO_2 concentration. In this context, the testimony Dr. Legates that an arbitrary increase in greenhouse gases would lead only to a "slight" increase in termperature, seems especially disingenuous. Dr. Legates seems to be suggesting that the warming would be small despite the magnitude of the CO_2 increase. Yet, both model-based studies and analyses of how climates changes in the past may have been influenced by changes in atmospheric composition suggest that a 1.5° C to 4.5° C increase in temperature is likely for each doubling of the CO_2 concentration. Thus, a quadrupling of the CO_2 concentration, which is plausible if the world chooses to derive most of its future energy from coal, would be expected to be associated with a roughly 3° C to 9° C increase in global mean temperature. Does Dr, Legates consider this a "slightly" increased temperature?

Question 29. In his testimony, Dr. Legates indicated that there were historical cases where the temperature has gone up; but the CO_2 has fallen. Do you agree there were such periods and how would you explain this?

Response. It is certainly the case that this has happened in the past. However, it is hardly surprising, and certainly not inconsistent with our established understanding of the various factors that influence surface temperatures. The warming response to increased greenhouse gas concentrations lags the actual increase in greenhouse gas concentrations in the atmosphere potentially by several decades, due to the sluggish response of the oceans, which have an enormous thermal capacity compared to the atmosphere, to increased surface radiative forcing. So warming is not expected to be contemporaneous with changes in CO_2, but instead, to lag it by several decades. In addition, greenhouse gases are certainly not the only factor affecting the average surface temperature of the Earth. There are other anthropogenic factors, such as increased sulphate aerosols, which can have a cooling effect on the climate, and natural factors, such as volcanic activity, modest natural variations in solar output, and internal dynamics associated with climate events such as El Niño, which also influence the average surface temperature of the globe. At any particular time, these other factors may outweigh the warming effect due to increased greenhouse gases. For example, the relative lack of warming during the period 1940–1970 appears to be related to a combination of such factors, as discussed in my response to an earlier question. But while these other factors tend to cancel over time, the increased greenhouse gases lead to a systematic warming that will not cancel out over time. It is for precisely this reason that late 20th century warming now appears to have risen above the range of the natural variability of past centuries.

There are two myths commonly perpetuated by climate change contrarians with regard to the relationship between historical CO_2 and temperature variations that are worth addressing in particular:

(1) Contrarians sometimes argue that the fact that the seasonal cycle in atmospheric CO_2, which is opposite of the seasonal cycle in temperature in the Northern Hemisphere (maximum atmospheric CO_2 levels over the course of the year occur during the Northern Hemisphere winter) implies a negative feedback of temperature on CO_2 concentration. Such an argument is based on a most profound misunderstanding of the basic principles governing atmospheric chemistry. Properly trained atmospheric chemists know that the seasonal cycle in global atmospheric CO_2 concentration is governed by the breathing of the terrestrial biosphere, which exhibits a hemispheric (and thus seasonal) asymmetry: there is a net uptake of atmospheric CO_2 (and thus a drawdown of atmospheric CO_2 concentrations) by terrestrial plants during the Northern Hemisphere summer growing season, owing to the vastly greater proportion of land in the Northern Hemisphere. This simple fact, and nothing else, dictates the relationship between Northern Hemisphere surface temperatures and CO_2 concentrations on seasonal timescales.

(2) Contrarians sometimes argue that the relationship between atmospheric CO_2 concentrations and temperature variations associated with glacial/interglacial variations over the past several hundred thousand years, as deduced from ice core

measurements shows that CO_2 is an effect, rather than cause, of climate variability. This reasoning is unsound for at least two fundamental reasons:

(a) Detailed measurements show that global atmospheric CO_2 concentrations lead estimated polar temperature variations (as deduced from ice core oxygen isotope ratios) during the long phase of increased glaciation, consistent with greenhouse gas forcing of the atmosphere. There is some evidence that CO_2 concentrations, however, lag estimated polar temperature variations during the rapid phase of deglaciation (melting of the terrestrial ice sheets at the termination of an ice age). This observation is the basis of the flawed argument summarized below. During this more rapid 'deglacial' phase, the climate system if far from being in an equilibrium state, and the dynamics of the climate system must be considered as representative of the a coupled interaction between surface temperature, atmospheric CO_2, ocean circulation, and glacial mass. It is well known by glaciologists who study this problem that the relationship between CO_2 and temperatures in such a scenario cannot be interpreted in terms of a simple causeeffect relationship.

(b) Even during the rapid deglaciation, the oxygen isotopes from the ice cores only provide an estimate of surface temperature variations in the proximity of the ice core (and a very imperfect one, owing to possible seasonal deposition biases and non-temperature influences on isotope fractionation). They certainly do not provide an estimate of hemispheric, let alone global, temperature variations. Thus, a comparison of ice core estimates of CO_2 and oxygen o18 ratios cannot be used to confidently infer the relationship between CO_2 concentrations and global mean temperatures

Question 30. During the hearing, there was some contention over the issue of the effect of surface cover changes and urban influences on the climate? Could you please restate your position on the likely sign and magnitude of the influence of both factors?

Response. Unfortunately misleading comments by Soon and Legates, and the complexity of the issues involved, made it difficult for me to convey, in the brief time allotted, the established science dealing with the various influences on Earth's surface radiation balance and changes therein in recent decades. Legates in his testimony confused and misstated the nature of both natural and anthropogenic influences on the Earth's surface energy budget and on the measurement of surface temperatures from surface-based stations. There are several different issues involved here, which I will attempt to clarify one at a time below:

(1) The claim made by Legates that the location of thermometer measurements in urban centers biases estimates of global mean temperature from the available meteorological observations would be correct were this effect not already carefully accounted for. In particular, possible urban heat island effects on global temperature estimates have been studiously accounted for in estimates that have been produced for more than a decade, See e.g. the 2001 IPCC report. This is unrelated to the issue of the influence of land-use changes on the surface radiation budget, though Legates testimony blurs the distinction between the two issues:

(2) The implication by Legates that land-use changes (such as urbanization) are the dominant influence on changes in the absorptive properties of the Earth as a whole in recent decades is completely wrong for at least two reasons:

(a) The primary factor impacting changes in the absorption of solar insolation by the Earth's surface in modern decades is the decrease in reflective snow and ice cover due to the warming of the Earth's surface. This represents a well-known positive feedback (the 'ice-albedo' feedback) associated with global warming in which warming leads to melting of snow and ice, which decreases the reflective properties of the surface, increasing surface absorption of radiation, and: thus increased the surface temperatures themselves. This crucial positive feedback, which enhances the impact of greenhouse gas concentrations on the warming of the surface, is fully accounted for in the climate model simulations that I have referred to above and in my testimony.

(b) While urbanization, as suggested by Legates, may lead to increased absorption of solar insolation in some urban areas, this is the more minor of the human land use changes impacting climate. There are far more extensive regions of the Earth where other changes in land use, such as conversion of forested land to agricultural land, have, instead, increased the reflective properties of the Earth's surface [Ramankutty, N., and J. A. Foley Estimating historical changes in global land cover: croplands from 1700 to 992, Global Biogeochemical Cycles, 13, 997–1027, 1999.], tending to cool the surface, as I explained in my testimony. Scientists who have studied the influences of these effects have found that the latter cooling effect is the dominant of these two anthropogenic land-use influences on the Earth's surface properties. Thus, climate model simulations investigating the influence of land-use

changes on hemispheric or global mean temperatures indicate that they have imposed a modest cooling influence [Govindasamy, B., P.B. Duffy, and K. Caldeira, Land use changes and Northern Hemisphere cooling, *Geophysical Research Letters*, 28, 291–294, 2001; Bauer, E, M. Claussen, and V. Brovkin, Assessing climate forcings of the Earth system for the past millennium, *Geophys. Res. Lett., 30*, doi: 10.1029/2002GL016639, 2003] that partially offsets even greater warming that would have been realized during the 20th century due to anthropogenic greenhouse gas influences, Evidence, therefore, does not support the case, as argued by Legates, that the full range of human land use changes have had a net warming effect on the climate. They have had a modest cooling influence on the climate.

Question 31. Do you receive any income from any sources which have taken advocacy positions with respect to the Kyoto Protocol, the U.N. Framework Convention on Climate Change, or legislation before the U.S. Congress that would affect greenhouse gas emissions? If so, please identify those sources and the relevant advocacy position taken.

Response. I do not, nor have I ever, received any such income.

Statement of David R. Legates, Director, Center for Climatic Research, University of Delaware

Distinguished Senators, panelists, and members of the audience: I would like to thank the Committee for inviting my commentary on this important topic of climate history and its implications. My name is David R. Legates and I am an Associate Professor and Director of the Center for Climatic Research at the University of Delaware in Newark, Delaware. My research interests have focused on hydroclimatology—the study of water in the atmosphere and on the land—and on the application of statistical methods in climatological research.

I am familiar with the testimony presented here by Dr. Soon. I agree with his statements and I will not reiterate his arguments. My contributions to Dr. Soon's research stemmed from my grappling with the apparent technology between the long-standing historical record and the time-series recently presented by Dr. Mann and his colleagues. It also stems from my own experiences in compiling and merging global estimates of air temperature and precipitation from a variety of disparate sources.

My Ph.D. dissertation resulted in the compilation of high-resolution climatologies of global air temperature and precipitation. From that experience, I have become acutely aware of the issues associated with merging data from a variety of sources and containing various biases and uncertainties. By its very nature, climatological data exhibit a number of spatial and temporal biases that must be taken into account. Instrumental records exist only for the last century or so and thus proxy records can only be used to glean information about the climate for earlier time periods. But it must be noted that proxy records are *not* observations and strong caveats must be considered when they are used. It too must be noted that observational data are not without bias either.

THE HISTORICAL RECORD OF THE LAST MILLENNIUM

Much research has described both the written and oral histories of the climate as well as the proxy climate records (e.g., ice cores, tree rings, and sedimentations) that have been derived for the last millennium. It is recognized that such records are not without their biases—for example, historical accounts often are tainted with the preconceived beliefs and limited experiences of explorers and historians while trees and vegetation respond not just to air temperature fluctuations, but to the entire hydrologic cycle of water supply (precipitation) and demand (which is, in part, driven by air temperature). Nevertheless, such accounts indicate that the climate of the last millennium has been characterized by considerable variability and that extended periods of cold and warmth existed. It has been generally agreed that during the early periods of the last millennium, air temperatures were warmer and that temperatures became cooler toward the middle of the millennium. This gave rise to the terms the "Medieval Warm Period" and the "Little Ice Age", respectively. However, as these periods were not always consistently warm or cold nor were the extremes geographically commensurate in time, such terms must be used with care.

A BIASED RECORD PRESENTED BY THE IPCC AND NATIONAL ASSESSMENT

In a change from its earlier report, however, the Third Assessment Report of the Intergovernmental Panel on Climate Change (IPCC), and now the U.S. National Assessment of Climate Change, both indicate that hemispheric or global air tempera-

tures followed a curve developed by Dr. Mann and his colleagues in 1999. This curve exhibits two notable features. First is a relatively flat and somewhat decreasing trend in air temperature that extends from 1000AD to about 1900AD and is associated with a relatively high degree of uncertainty. This is followed by an abrupt rise in air temperature during the 1900's that culminates in 1998 with the highest temperature on the graph. Virtually no uncertainty is shown for the data of the last century. The conclusion reached by the IPCC and the National Assessment is that the 1990's are the warmest decade with 1998 being the warmest year of the last millennium.

Despite the large uncertainty, the surprising lack of variability in the record gives the impression that climate remained relatively unchanged through most of the last millennium—at least until human influences began to cause an abrupt increase in temperatures during the last century. Interestingly, Mann *et al.* replace the proxy data for the 1900's by the instrumental record and no uncertainty characterization is provided. This too yields a false impression that the instrumental record is consistent with the proxy data and that it is 'error free'. It is neither. The instrumental record contains numerous uncertainties, resulting from a lack of coverage over the world's oceans, an under-representation of mountainous and polar regions as well as under-developed nations, and the presence of urbanization effects resulting from the growth of cities. Even if a modest uncertainty of a ±0.1° C were imposed on the instrumental record, the claim of the 1990's being the warmest decade would immediately become questionable, as the uncertainty window would overlap with the uncertainty associated with earlier time periods. Note that if the satellite temperature record—where little warming has been observed over the last 20 years—had been inserted instead of the instrumental record, it would be impossible to argue that the 1990's are the warmest decade.

RATIONALE FOR THE SOON *et al.* investigation

So we were left to question why the Mann *et al.* curve seems to be at variance with the previous historical characterization of climatic variability. Investigating more than several hundred studies that have developed proxy records, we came to the conclusion that nearly all of these records show considerable fluctuations in air temperature over the last millennium. Please note that we did *not* reanalyze the proxy data—the original analysis from the various researchers was left intact. Most records show the coldest period is commensurate with at least a portion of what is termed the "Little Ice Age" and the warmest conditions are concomitant with at least a portion of what is termed the "Medieval Warm Period".

But our conclusion is entirely consistent with conclusions reached by Drs. Bradley and Jones that not all locations on the globe experienced cold or warm conditions simultaneously. Moreover, we chose *not* to append the instrumental record, but to compare apples-with-apples and determine if the proxy records *themselves* indeed confirm the claim of the 1990's being the warmest decade of the last millennium. That claim is *not* borne out by the individual proxy records.

However, the IPCC report, in the chapter with Dr. Mann as a lead author and his colleagues as contributing authors, also concludes that research by Drs. Mann, Jones, and their colleagues "support the idea that the 15th to 19th centuries were the coldest of the millennium over the Northern Hemisphere overall." Moreover, the IPCC report also concludes that the Mann and Jones research "show[s] temperatures from the 11th to 14th centuries to be about 0.2° C warmer than those from the 15th to 19th centuries." This again is entirely consistent with our findings. Where we differ with Dr. Mann and his colleagues is in their construction of the hemispheric averaged time-series, their assertion that the 1990's are the warmest decade of the last millennium, and that human influences appear to be the only significant factor on globally averaged air temperature. Reasons why the Mann *et al.* curve fails to retain the fidelity of the individual proxy records are detailed statistical issues into which I will not delve. But our real difference of opinion focuses solely on the Mann *et al.* curve and how we have concluded it misrepresents the individual proxy records. In a very real sense, this is an important issue that scientists must address *before* the Mann *et al.* curve is taken as fact.

Our work has been met with much consternation from a variety of sources and we welcome healthy scientific debate. After all, it is disagreements among scientists that often lead to new theories and discoveries. However, I am aware that the editors of the two journals that published the Soon *et al.* articles have been vilified and the discussion has even gone so far as to suggest that Drs. Soon and Baliunas be barred from publishing in the journal *Climate Research*. Such tactics have no place in scientific debate and they inhibit the free exchange of ideas that is the hallmark of scientific inquiry.

CLIMATE IS MORE THAN MEAN GLOBAL AIR TEMPERATURE

In closing, let me state that climate is more than simply annually-averaged global air temperature. Too much focus has been placed on divining air temperature time-series and such emphasis obscures the true issue in understanding climate change and variability. If we are truly to understand climate and its impacts and driving forces, we must push beyond the tendency to distill it to a single annual number. Proxy records, which provide our only possible link to the past, are incomplete at best. But when these records are carefully and individually examined, one reaches the conclusion that climate variability has been a natural occurrence, and especially so over the last millennium. And given the uncertainties in the proxy and instrumental records, an assertion of any decade as being the warmest in the last millennium is premature.

I'm sorry that a discussion that is best conducted among scientists has made its way to a U.S. Senate committee. But hopefully a healthy scientific debate will not be compromised and we can push on toward a better understanding of climate change.

I again thank you for the privilege of *speaking* before you today.

PREPARED STATEMENT OF LEONARD LEVIN, PH.D., TECHNICAL LEADER, EPRI, PALO ALTO, CALIFORNIA

I am Leonard Levin, technical leader at EPRI, which is a non-profit, collaborative organization conducting energy-related R&D in the public interest. Our members are public and private organizations in the electricity and energy fields, and we now serve more than 1000 energy and governmental organizations in more than 40 countries. These remarks constitute a synthesis of current research on environmental mercury, and are not a representation of official EPRI position.

INTRODUCTION

As a global pollutant, the impact of mercury on the human environment is a critical issue that EPRI and the scientific community have been examining for many years. As the scientific understanding of where mercury originates nationally and globally, combined with the new health data, continues to be refined, it can help inform decisions regarding its management. I would like to address three key questions where new findings have emerged. First, where does mercury found in the U.S. environment originate? Second, how much has mercury in fish changed in the last few decades? And third, how do potential mercury management steps change the amount of mercury depositing to the earth's surface in the U.S.?

WHERE DOES MERCURY IN THE U.S. ENVIRONMENT ORIGINATE?

Mercury is clearly a global issue. Recent estimates are that, in 1998, some 2340 tons of mercury were emitted globally through industrial activity; of these, more than half, or 1230 tons, came from Asian countries, primarily China[1]. These findings are similar to those of other global inventories[2]. In addition, it is estimated that another 1300 tons of mercury emanates from land-based natural sources globally, including abandoned mining sites and exposed geological formations. Another 1100 tons or so issues from the world's oceans, representing both new mercury emitted by undersea vents and volcanoes, and mercury cycled through the ocean from the atmosphere previously. Recent findings from the large United States-Canadian METAALICUS field study in Ontario, Canada showed that a fairly small amount of deposited mercury, no more than 20 percent or so, re-emits to the atmosphere, even over a 2-year period. The implications of this are profound: mercury may be less mobile in the environment than we previously thought; once it is removed from the atmosphere, it may play less of a role in the so-called "grasshopper effect[3]" where persistent global pollutants are believed to successively deposit and re-emit for many years and over thousands of miles.

---

[1] Christian Seigneur, C., K. Vijayaraghavan, K. Lohman, Pr. Karamchandani, C. Scott, Global Source Attribution for Mercury Deposition in the United States, submitted to Environ. Sci. Technol., 2003.

[2] Jozef M. Pacyna, Elisabeth G. Pacyna, Frits Steenhuisen and Simon Wilson, Mapping 1995 global anthropogenic emissions of mercury, Atmospheric Environment 37 (S1) (2003) pp. 109–117.

[3] Environment Canada, The Grasshopper Effect and Tracking Hazardous Air Pollutants, The Science and the Environment Bulletin, May/June 1998.

Recent studies by EPRI have shown that the mercury depositing into the U.S. from the atmosphere may originate at very distant points. Model assessments show that, for ¾ of the area of the continental United States, more than 60 percent of the mercury received originates outside U.S. borders, from other countries or even other continents. Only 8 percent of U.S. territory receives ⅔ or more of its mercury from U.S. domestic sources, and less than 1 percent of U.S. territory gets 80 percent or more of its mercury from sources within the U.S. One implication of this dichotomy between mercury sources and the U.S. areas impacted is that there may be a "management floor" for U.S. mercury, a level below which the amount of mercury depositing to the surface cannot be reduced.

Additional evidence for the external origins of much of the mercury in the U.S. environment was gathered over the last 2 years by aircraft experiments carried out by EPRI, the National Center for Atmospheric Research, and a number of U.S., Asian, and Australian investigators. One set of flights measured significant mercury in winds entering the Pacific Ocean from Shanghai; China; researchers tracked the Chinese mercury plume over the Pacific for 400 miles toward America. A second set of flights from Monterey, California, found that same plume from China crossing the California coast, and a second, higher plume of enriched mercury originating in Central Asia also moving into the U.S. The global nature of mercury in the U.S. has been clearly demonstrated.

WHAT ARE THE PRIMARY SOURCES OF MERCURY IN FISH AND THE ENVIRONMENT?

For much of the twentieth century, mercury was an essential part of industrial products, such as batteries and switches, or a key ingredient in such other products as house paints. These industrial uses of the element declined significantly in the latter half of the century, and are now less than 10 percent of their use of fifty years ago.[4] Professor Francois Morel of Princeton University and colleagues recently analyzed newly caught Pacific tuna for mercury[5], and compared those results to the mercury content of similar tuna caught in the 1970's. Despite changes in mercury emissions to the atmosphere in those thirty years[6], and a matching increase in the mercury depositing; from the atmosphere to rivers and oceans below, Prof Morel found that mercury levels in tuna have not changed over that time. One conclusion is that the mercury taken up by such marine fish as tuna is not coming from sources on land, such as utility power plants, but from natural submarine sources of mercury, including deep sea volcanoes and ocean floor vents. The implications are that changes in mercury sources on the continents will not affect the mercury levels found in open ocean foodfish like tuna.

An estimate in 2001 by scientists of the Geological Survey of Canada and others[7] estimated that geological emissions of mercury, as well as emissions from inactive industrial sites on land, are five to seven times as large as had been estimated earlier. Recent measurements in the stratosphere by EPRI researchers show a rapid removal of mercury in the upper atmosphere, allowing for additional sources at the surface while still maintaining the measured rates of deposition and removal needed for a global balance of sources and sinks. As a result, it is now possible to attribute a greater fraction of the mercury entering U.S. waters to background natural sources rather than industrial emissions from the U.S. or elsewhere globally.

HOW COULD POTENTIAL MERCURY REDUCTIONS CHANGE MERCURY DEPOSITION?

EPRI recently completed work to assess the consequences, of further mercury emissions reductions for mercury in the atmosphere, U.S. waterways, and fish[8]. The approach used linked models of atmospheric mercury chemistry and physics with analyses of Federal data on mercury in fish in the U.S. diet, along with a model of costs needed to attain a given reduction level.

[4] Engstrom, D.R., E.B. Swain, Recent Declines in Atmospheric Mercury Deposition in the Upper Midwest, Environ. Sci. Technol. 1997, 31, 960–967.

[5] Kraepiel, A.M.L., K. Keller, H.B. Chin, E.G. Malcolm, F.M.M. Morel, Sources and Variations of Mercury in Tuna, Meeting of American Society for Limnology and Oceanography, Salt Lake City, Utah: January 2003.

[6] Slemr, F., E-G. Brunke, R. Ebinghaus, C. Temme, J. Munthe, I. Wangberg, W. Schroeder, A. Stgeffen, T. Berg, Worldwide trend of atmospheric mercury since 1977, Geophys. Res. Ltrs., 30 (10), 23–1 to 23–4

[7] Richardson G. M., R. Garrett, I. Mitchell, M. Mah-Paulson, T. Hackbarth, Critical Review On Natural Global And Regional Emissions Of Six Trace Metals To The Atmosphere, International Lead Zinc Research Organization, International Copper Association, Nickel Producers Environmental Research Association.

[8] EPRI Technical Report 1005224, "A Framework for Assessing the Cost-Effectiveness of Electric Power Sector Mercury Control Policies," EPRI, Palo Alto, CA, May 2003.

Current U.S. utility emissions of mercury are about 46 tons per year. At the same time, a total of about 179 tons of mercury deposit each year in the U.S., from all sources global and domestic. One proposed management scenario examined cutting these utility emissions by 47 percent, to 24 tons per year. The analysis showed that this cut results in an average 3 percent drop in mercury deposition into the U.S. Some isolated areas totaling about 1 percent of U.S. land area experience drops of up to 30 percent in mercury deposited. The cost model used in association with these calculations showed utility costs to reach these emission control levels would amount to between $2 billion and $5 billion per year over 12 years. This demonstrated that U.S. mercury patterns are relatively insensitive to the effects of this single category of sources.

In addition, most of the fish consumed in the U.S. comes from ocean sources, which would be only marginally affected by a global reduction of 24 tons of mercury per year due solely to U.S. controls. Wild fresh water fish in the U.S. would be expected to show a greater reduction in mercury content, but are a relatively small part of the U.S. diet compared to ocean or farmed fish. When these changes were translated into how much less mercury enters the U.S. diet, we found that 0.064 percent fewer children would be born "at risk" due to their mothers taking in less mercury from consumed fish. These results were based on the Federal dietary fish consumption data. So, a drop of nearly half in utility mercury emissions results in a drop of 3 percent (on average) in mercury depositing to the ground, and a drop of less than one-tenth of a percent in the number of children "at risk."

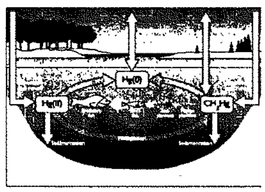
DECISIONMAKING UNDER UNCERTAINTY

These recent findings on mercury sources, dynamics, and management are a small part of the massive international research effort to understand mercury and its impacts on the human environment. EPRI and others, including the U.S. Environmental Protection Agency and the U.S. Department of Energy, are racing to clarify the complex interactions of mercury with geochemical and biological systems, vital to understanding mercury's route to human exposure and potential health effects. With this improved understanding, informed decisions can be made on the best ways to manage mercury.

Thank you for the opportunity to deliver these remarks to the Committee.

research brief

Framework Developed for Assessing Mercury Control Policies



An analysis of alternative proposed mercury emission control strategies reveals substantial differences in the costs and effectiveness of contrasting regulatory approaches to achieving emissions reductions, according to the results of a new study published by EPRI.

The study, an initial exploration of the costs associated with potential limits on mercury emissions from U.S. coal-fired power plants and the effectiveness of those limits in reducing exposure to mercury (Hg) among sensitive groups, assesses emissions scenarios for electric utilities under both the Clear Skies Act (CSA) of 2002, introduced in the U.S. Congress in 2003 in substantially the same form, and an illustrative, maximum achievable control technology (MACT) standard.

An EPRI research team developed a framework to incorporate current Hg research results into an economic model able to assess potential policy. The framework incorporates these findings under alternative regulatory regimes and estimates changes in mercury emission, by mercury species, to calculate regional changes in mercury deposition. These analyses were then used to estimate the resulting reduction in human intake of methylmercury via fish consumption, compared to current regulatory health standards.

Common framework

One basis for comparing proposed regulatory steps is a common cost-effectiveness framework. Such a framework can assist decision-making regarding expenditures to reduce Hg deposition, including the change in co-benefits of capturing mercury by controlling other emissions under different scenarios, costs versus exposure, or other cost-effectiveness measures. The current project reflects the multivariate nature of mercury management considerations, from changes in atmospheric loading and deposition to changes in health impacts on society, and will assist in informing development of public policy.

Analytic approach

The project team first examined the utility sector's response to mercury control policies potentially applied to coal-fired electricity generating plants. The team used the Electric Power Market Model (EPMM) to project the costs and emission changes expected from regulations, including the effects of any permitted mercury emission trading. EPMM determines the most efficient (least-cost) manner to satisfy regional electric loads in twenty-eight continental U.S. regions until year 2020.

The team modeled the environmental fate and transport of Hg emissions, using emission factors for mercury speciated into its two key valence states: elemental and divalent. The proportion of each species emitted is highly dependent on the type of coal burned and emission controls installed on each plant. To estimate the speciation of Hg within various plant types, algorithms were developed and applied based on EPRI's analysis of U.S. Environmental Protection Agency information collection request (ICR) data on mercury in coal burned by utilities.

Deposition changes between cases and various control scenarios were estimated using source-receptor relationships that Atmospheric and Environmental Research (AER) developed for EPRI with its Trace Element Analysis Model (TEAM). The regional speciated emissions from EPMM were used with this source-receptor model to estimate the change in Hg deposited in nineteen receptor regions across the United States.

Data on fish consumption of U.S. women with high-end blood levels of mercury were analyzed, and estimates were made of the fraction of consumption of farmed fish, which are assumed to be insensitive to mercury changes due to controlled diets. The change in content of MeHg in wild freshwater fish was assumed to vary linearly with the change in deposition. This deposition change was used to estimate the change in MeHg exposure to women of childbearing age from consumption of wild freshwater fish. Reductions in MeHg in marine fish from the global ocean also were included in the overall exposure reduction analysis.

Key assumptions

Several assumptions were made within the framework. Foremost, it was assumed that all MeHg exposure occurs

through fish consumption, as was assumed in EPA's 1997 *Mercury Study Report to Congress*. Second, the findings conservatively assumed any reductions in Hg deposition will immediately translate into comparable reductions in MeHg concentration in fish, ignoring the lag time known to be required to reduce Hg concentration in water, sediments, and fish tissue. Third, the effectiveness of a policy was evaluated in terms of the reduced number of births in which maternal blood MeHg exceeds the EPA reference dose (RfD). In addition, the reduction in MeHg exposure to all women, on average, is calculated for various mercury control policies.

Study results

By 2020, CSA's present-value incremental cost is estimated to be approximately $6 billion, while MACT's present value incremental cost is approximately $19.3 billion. Although much less expensive, CSA ultimately produces greater mercury reductions than the illustrative MACT.

In 2020, CSA reduces continental U.S. Hg deposition by an average of 1.5%, reduces exposure of women of childbearing age to methylmercury by 0.5%, and decreases the fraction of the population above the RfD by 0.064%. In comparison, the illustrative MACT reduces deposition by 1.2% and decreases the population fraction above the RfD by 0.055% by 2020. A primary difference between the two scenarios is that the MACT brings about Hg exposure changes more quickly.

EPRI perspective

"Informed public policy requires a comprehensive and logically consistent technical assessment of the costs and effectiveness of alternative courses of action," explains Leonard Levin, EPRI technical leader and program manager for air toxics. "The new report describes one of the first efforts to integrate available research and analysis capabilities related to mercury in the environment to better inform decision-makers."

To obtain a copy of the new report, *A Framework for Assessing the Cost-Effectiveness of Electric Power Sector Mercury Control Policies* (Report 1005224), call EPRI Customer Service, (800) 313-3773. *For more information, contact Leonard Levin, llevin@epri.com, (650) 855-7929.*

EPRI

A Framework for Assessing the Cost-Effectiveness of Electric Power Sector Mercury Control Policies

Technical Report

A Framework for Assessing the Cost-Effectiveness of Electric Power Sector Mercury Control Policies

1005224

Final Report, May 2003

EPRI Project Manager
L. Levin

EPRI • 3412 Hillview Avenue, Palo Alto, California 94304 • PO Box 10412, Palo Alto, California 94303 • USA
800.313.3774 • 650.855.2121 • askepri@epri.com • www.epri.com

DISCLAIMER OF WARRANTIES AND LIMITATION OF LIABILITIES

ORGANIZATION(S) THAT PREPARED THIS DOCUMENT

Charles River Associates, Incorporated

ENVIRON International Corporation

ORDERING INFORMATION

Requests for copies of this report should be directed to EPRI Orders and Conferences, 1355 Willow Way, Suite 278, Concord, CA 94520, (800) 313-3774, press 2 or internally x5379, (925) 609-9169, (925) 609-1310 (fax).

CITATIONS

This report was prepared by

Charles River Associates, Incorporated
1201 F Street, N.W., Suite 700
Washington D.C. 20004-1204

Principal Investigators
A. Smith
J. Rego
B. Lonergan

ENVIRON International Corporation
6001 Shellmound Street, Suite 700
Emeryville, CA 94608

Principal Investigator
C. Whipple

This report describes research sponsored by EPRI.

The report is a corporate document that should be cited in the literature in the following manner:

A Framework for Assessing the Cost-Effectiveness of Electric Power Sector Mercury Control Policies, EPRI, Palo Alto, CA: 2003. 1005224.

REPORT SUMMARY

This report is an initial exploration of the costs associated with potential limits on mercury emissions from U.S. coal-fired power plants and the effectiveness of those limits in reducing exposure to mercury among sensitive groups. The report assesses mercury emissions scenarios for electric utilities under two alternative publicly proposed Hg control policies: the Clear Skies Act (CSA) of 2002 and an illustrative maximum achievable control technology (MACT) standard.

The project team developed a framework to incorporate current Hg research data into an economic-based design able to assess potential policy legislation. The framework incorporates these data under alternative regulatory regimes and estimates changes in mercury emission, by mercury species, to calculate regional changes in mercury deposition. These analyses are used to estimate the resulting reduction in intake of methylmercury by humans via fish consumption.

Results & Findings

By 2020, CSA's present-value incremental cost is estimated to be approximately \$6 billion; MACT's present value incremental cost is approximately \$19.3 billion. Although much less expensive, CSA ultimately produces greater mercury reductions than the illustrative MACT. In 2020, CSA reduces continental U.S. Hg deposition by an average of 1.5%, reduces exposure of women of childbearing age to methylmercury (MeHg) by 0.5%, and decreases the fraction of the population above the RfD by 0.064%. In comparison, the illustrative MACT reduces deposition by 1.2% and decreases the population fraction above the RfD by 0.055% by 2020. The other primary difference between the two scenarios is that the MACT brings about Hg exposure changes more quickly.

Challenges & Objectives

One basis for comparing proposed regulatory steps is a common cost-effectiveness framework. Such a framework can assist decision-making regarding expenditures to reduce Hg deposition, including the change in co-benefits under different scenarios, costs versus exposure, or other cost-effectiveness measures. This project reflects the multivariate nature of mercury management considerations, from changes in atmospheric loading and deposition to changes in health impacts on society, and will assist in informing development of public policy.

Applications, Values & Use

This report presents a framework for policy assessment; however, future refinements would help reach an optimal decision-making framework and, thus, an optimal policy. Due to data limitations concerning consumption patterns for wild freshwater fish, geographical differences in the cost-effectiveness of control requirements have not been incorporated in this analysis. Developing better information on MeHg and fish consumption patterns of the most at-risk portions of the U.S. population will allow future research to refine source-receptor-consumption relationships.

EPRI Perspective
Informed public policy requires a comprehensive and logically consistent technical assessment of the costs and effectiveness of alternative courses of action. This report describes one of the first efforts to integrate available research and analysis capabilities related to mercury in the environment to better inform decisionmakers.

Approach
The project team first examined the utility sector's response to coal-fired electricity generating plant mercury control policies. The team used the Electric Power Market Model (EPMM) to project the costs and emission changes expected from regulations, including the effects of permitted trading. EPMM determines the most efficient (least-cost) manner to satisfy regional loads in twenty-eight continental U.S. regions until year 2020. The team modeled fate and transport of Hg emissions, using emission factors for mercury speciated into its two key valence states: elemental (Hg^0) and ionic (Hg^{2+}). The proportion of each species emitted is highly dependent on the type of coal burned and emission controls installed on each plant. To estimate the speciation of Hg within various plant types, algorithms were developed and applied based on EPRI's analysis of EPA's information collection request (ICR) data.

Deposition changes were estimated using source-receptor relationships that Atmospheric and Environmental Research (AER) developed with their Trace Element Analysis Model (TEAM). The regional speciated emissions from EPMM were used with this source-receptor model to estimate the change in Hg deposited in nineteen receptor regions across the United States. The change in content of MeHg in wild freshwater fish was assumed to vary linearly with the change in deposition. This deposition change was used to estimate the change in MeHg exposure to women of childbearing age from consumption of wild freshwater fish. Reductions in MeHg in marine fish from the global ocean also were included in the overall exposure reduction analysis.

Several assumptions were made within this framework. Foremost, it was assumed that all MeHg exposure occurs through fish consumption, as was assumed in EPA's 1999 *Mercury Study Report to Congress*. Second, the findings conservatively assumed any reductions in Hg deposition will immediately translate into comparable reductions in MeHg concentration in fish, ignoring the lag time naturally required to reduce Hg concentration in water, sediments, and fish tissue. Third, the effectiveness of a policy was evaluated in terms of the reduced number of births in which maternal blood MeHg exceeds the EPA reference dose (RfD). In addition, the reduction in MeHg exposure to all women, on average, is calculated for various mercury control policies.

Keywords
Mercury
Methylmercury
Cost-effectiveness
Control policy
Clear skies act
MACT
Multi-pollutant legislation

CONTENTS

1 INTRODUCTION 1-1

2 COST-EFFECTIVENESS FRAMEWORK AND DATA SOURCES 2-1

2.1 Electricity Generating System Operation 2-1

2.2 Hg Speciation 2-3

2.3 Deposition Changes 2-4

2.4 Changes in Human Mercury Exposure 2-7

2.5 Characteristics of Methylmercury Exposures and the Effect of Exposure Reductions on the Number at Risk 2-10

2.6 Cost-Effectiveness: The Final Step 2-12

3 ASSESSMENT OF TWO ALTERNATIVE MERCURY CONTROL POLICIES 3-1

3.1 Clear Skies Act 3-2

3.2 Illustrative MACT 3-9

3.3 Comparing Control Policy Options 3-11

4 CONCLUSION 4-1

5 REFERENCES 5-1

A DESCRIPTION OF THE EPMM MODEL A-1

A.1 Overall Analytical Approach A-1

A.2 The Electric Power Market Model (EPMM) A-4

A.3 Projecting Competitive Energy and Capacity Prices A-6

A.4 Other Impact Projections A-8

A.5 Key Assumptions A-8

A.6 Data Sources A-10

B **DETAILS OF CALCULATIONS TO SPECIATE EPMM'S TOTAL HG EMISSIONS PROJECTIONS B-1**

C **DETAILS OF RECEPTOR LOCATIONS FOR TRANSFER COEFFICIENTS BASED ON AER'S TEAM MODEL C-1**

D **BLOOD LEVELS OF METHYLMERCURY AND FISH CONSUMPTION: INFORMATION FROM THE NHANES 2002 DATA RELEASE D-1**

D.1 Introduction D-1
D.2 Blood Concentrations D-1
D.3 Fish Consumption D-4
D.4 Fish Consumption versus Mercury Concentration D-4
D.5 Summary D-8

LIST OF FIGURES

Figure 2-1 Computational Elements of Hg Cost-Effectiveness Framework.............................. 2-2
Figure 2-2 Transfer Coefficients for Four Receptor Locations 2-5
Figure 2-3 Regression on NHANES Data to Estimate Slope of Distribution Near RfD............2-11
Figure 3-1 Hg^0 and Hg^{2+} Emissions Projections – Base Case 3-5
Figure 3-2 Hg^0 and Hg^{2+} Emissions Projections – CSA Case... 3-6
Figure 3-3 Hg^0 and Hg^{2+} Emissions Projections – MACT Case 3-10
Figure 3-4 Cost per Birth Moved Below the RfD under CSA and MACT................................3-13
Figure 3-5 Regional Deposition Change Comparison for CSA vs. MACT, Year 2020.............3-14
Figure A-1 Integrated Framework for Multi-Pollutant Control Policy Analysis A-2
Figure A-2 Regional Detail in EPMM.. A-5
Figure D-1 Distribution of Top 10% of MeHg Blood Levels, U.S. Women, Ages 16-49, with Demographic Weights Applied D-3
Figure D-2 Blood Mercury Versus Fish Meals Women Ages 16-49..D-6

LIST OF TABLES

Table 3-1 Differences in Coal Consumption in Trillions of Btu (CSA Compared to Base) 3-2
Table 3-2 Capacity (GW) Retrofitted in Base Scenario 3-3
Table 3-3 Capacity (GW) Retrofitted in CSA Scenario 3-4
Table 3-4 CSA's Incremental Cost Relative to Base Case ($millions in 1999 Dollars) 3-5
Table 3-5 Domestic Deposition Changes by Receptor Site (CSA compared to Base) 3-7
Table 3-6 National Average Cost-Effectiveness Summary for CSA Scenario (All Changes Are Relative to the Base Scenario Outcomes) 3-9
Table 3-7 Capacity (GW) Retrofitted in MACT Scenario 3-10
Table 3-8 National Average Cost-Effectiveness Summary for MACT Scenario (All Changes Are Relative to the Base Scenario Outcomes) 3-11
Table 3-9 Domestic Deposition Changes by Receptor Site (MACT compared to Base) 3-12
Table A-1 Examples of Integrated Framework Outputs A-3
Table B-1 "Industry Consensus" on Hg Co-Control Assumptions B-2
Table B-2 Hg Emissions Speciation Assumptions B-4
Table D-1 Blood Mercury Levels of Women Ages 16–49 .... D-2
Table D-2 Fish Consumption (Meals in the Previous 30 Days) Versus Blood Mercury Levels (Unweighted for Demographic Factors) D-5
Table D-3 Fish Meals Reported for the Previous 30 Days by Women Ages 16–49 and Children Ages 1-5 .... D-7
Table D-4 Summary of NHANES Data on Fish Meals D-8

1 INTRODUCTION

In December 2000, the U.S. Environmental Protection Agency (EPA) announced its determination that mercury (Hg) emissions from coal-fired steam electricity generating units should be subject to control as a toxic air pollutant. EPA also indicated that, as part of its rulemaking process, it would act under the maximum achievable control technology (MACT) rules. One issue still to be resolved is an approach for comparing the costs of such a MACT rule with the effectiveness of the reductions in mercury emissions that would be experienced under it. Since mercury is a multimedia pollutant, with little of the exposure of concern occurring by inhalation but most taking place by ingestion of mercury-containing fish, the effectiveness of controls will have to be gauged by changes in human exposure. Other potential avenues for regulating Hg emissions from electricity generation have been proposed. Various bills that have been introduced in the U.S. Congress share the feature that they would implement new emissions control programs for NO_x and SO_2 as well as Hg. (Some of the bills also include carbon constraints.) These bills usually rely on cap-and-trade programs for the emissions controls, often including Hg trading.

In this paper, we describe an integrated analysis framework that we have developed to assess the cost-effectiveness of Hg controls. The framework integrates information from a variety of sources:

- A model of electric utility responses to required controls, the Electric Power Market Model (EPMM), which projects the costs and emissions changes of electric sector regulations.
- The EPRI analysis of EPA's 1999-2000 utility mercury Information Collection Request (ICR) data, which permit us to estimate the effect of various technologies on Hg emissions speciation.
- The Trace Element Analysis Model (TEAM) of global mercury fate and transport developed by Atmospheric and Environmental Research (AER).
- The 1999 National Health and Nutrition Examination Survey (NHANES) data on fish consumption and blood methylmercury (MeHg) levels.
- The Census Bureau's national population projections.

The framework envisages emissions of mercury from electricity generation under alternative regulatory regimes at a regional spatial scale, and by mercury species (valence state). It then tracks these emissions through the mercury cycle to uptake by humans via consumption of mercury-containing fish. The framework's outputs include measures of effect such as changes in the total cost of producing electricity, changes in speciated mercury emissions, changes in deposition, and the impact of altered environmental mercury levels on exposures of women in the upper-end of the MeHg blood-level distribution. The last of these is used to estimate

the change in the number of births per year to women with MeHg blood levels above the EPA-determined reference dose (RfD) of 0.1 µg/kg-d. This measure is not meant to quantify total benefits, but to indicate the size of the MeHg blood level distribution shift for policy comparison purposes. Using any or all the foregoing measures, one can compare the relative cost-effectiveness of alternative MACT standards, and other policies for controlling utility Hg emissions.

The remainder of this paper is organized as follows. In Section 2 we explain the Hg cost-effectiveness framework in detail and describe the data sources used in constructing the framework. In Section 3 we apply the framework to compare the cost-effectiveness of the mercury limits specified in the Clear Skies Act with the cost-effectiveness of one possible MACT standard. Section 4 discusses possible framework refinements and extensions.

2
COST-EFFECTIVENESS FRAMEWORK AND DATA SOURCES

Figure 2-1 depicts the key components of the framework and the connections among them. The basic outputs of the framework are the incremental costs and incremental changes in human health risk from one regulatory scenario relative to another. "Cost-effectiveness" can be assessed by comparing any of the incremental impact estimates to the incremental costs of the policy scenario. For most of our summaries of policy options, we define "effectiveness" as the reduced number of births in which maternal blood MeHg exceeds the EPA RfD. However, other metrics could be used; for example, the information provided in this assessment also indicates the percent reduction in blood MeHg levels for all women. To compare among many different regulatory options, one would usually estimate the cost-effectiveness of each option relative to the same "Base" scenario. The Base scenario could be any arbitrary regulatory future, but optimally would be one that has no explicit Hg emissions controls.

As Figure 2-1 shows, estimation of the incremental human health risk changes requires a sequence of linked computational steps. Starting from estimated emissions changes, the framework estimates the extent to which the different Hg species are being changed, then converts these speciated emissions changes into projected deposition changes, which are in turn used to estimate changes in human MeHg consumption, and hence changes in the number of humans that are "at risk." The emissions and deposition steps are all performed at a regional level of detail. The calculations and data sources for each of these steps are described in the rest of this section.

2.1 Electricity Generating System Operation

The first step of cost-effectiveness analysis for any policy is to project the costs and emissions changes expected to result from that particular set of electricity generation regulations. Both of these projections are simultaneously generated by an electric generation market model designed for multi-pollutant emissions policy analysis, making such a model a key component of the framework. We use the Electric Power Market Model (EPMM) for this purpose. Appendix A provides a detailed discussion of EPMM, the most relevant parts of which are summarized below.

EPMM simulates the operation of every electricity generating unit in the United States for each modeled future year. EPMM determines the least-cost way the aggregate set of units can satisfy their regional load and reserve margin requirements without exceeding specified limits (if any) on emissions of SO_2, NO_x, Hg, and CO_2. System operational decisions include altering the utilization of existing capacity, building new capacity, retiring existing capacity, installing retrofits, and switching the fuel types of existing units. EPMM minimizes costs over all modeled years simultaneously (reflecting "forward-looking decision makers"), and accounts for synergies in controlling multiple pollutants. This makes it important to specify realistically how other

pollutants are being regulated in the "Base" scenario, and to maintain the same assumptions about regulation on those pollutants in the Hg scenario.

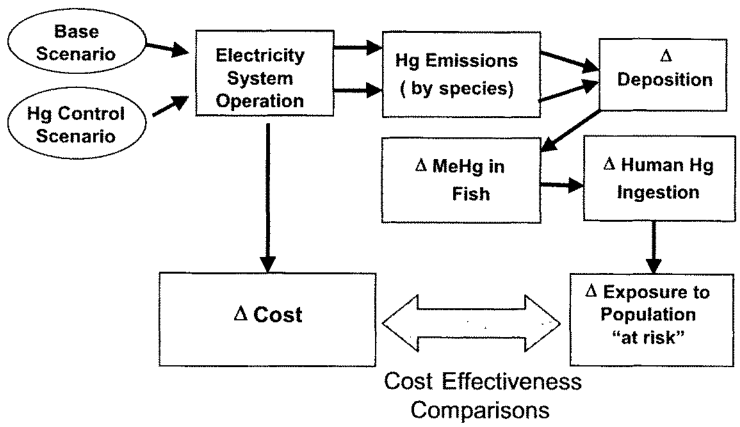


Figure 2-1
Computational Elements of Hg Cost-Effectiveness Framework

Note: The Delta (Δ) symbolizes a change between the Base and the Hg control policy scenario

Key inputs to EPMM include the set of electricity generating units and their characteristics (including emissions rates and pollution control equipment currently installed), fuel supply curves, regional electricity demand (load), limits on electricity transfers between regions, costs and removal rates of retrofit control options, and limits on emissions or emissions rates. The Hg emissions rates for each type of coal that might be burned are based on EPRI's analysis of the ICR data (EPRI, 2000).[1]

EPMM's outputs include projections of a large number of electricity generating system outcomes for 28 regions of the continental United States. Just a few key EPMM results are extracted for input into the cost-effectiveness framework: national costs by year and regional total Hg emissions by year. In addition, in order to disaggregate the Hg emissions to individual Hg species, the framework makes use of EPMM output on types of retrofits and types of coals being consumed in each region.[2]

[1] The EPMM model allows units to switch among a variety of coal types, and thus does not assume that units continue to burn the same coals that they reported in the ICR process. The switching decision is motivated by minimizing total cost per MWh, which accounts for the delivered coal prices plus any emissions costs. The ICR data were used to establish initial Hg emissions rates if utility plant operators do not switch coals; the Hg contents of coals that might be considered for switching to; and the amount of Hg removal that occurs with existing particulate and SO_2 controls.

[2] At present, EPMM produces only total Hg emissions estimates, and the speciation calculations are performed *ex post facto* in the cost-effectiveness framework. Eventually we plan to build the speciation calculations directly into the EPMM model, and then these additional EPMM outputs will no longer be needed by the framework.

The multi-pollutant capabilities of EPMM deserve some discussion. Some types of emissions control equipment designed to reduce SO_2 and NO_x may also reduce Hg emissions to some degree as a "co-control" benefit. Specifically, wet flue gas desulfurization (FGD) removes a large portion of the ionic Hg (because it is more water-soluble) before it can be emitted. Selective catalytic reduction (SCR), a leading NO_x control option, can help increase the proportion of the Hg that is in the ionic form, which improves its likelihood of removal by the FGD, if available.[3] As a result, even if there are no explicit Hg limits, one can expect some incremental reduction of Hg if SO_2 or NO_x caps are lowered. Also, significant limits or taxes on CO_2 emissions will lead to substitution from coal-fired generation to gas-fired generation, which reduces Hg emissions. Thus, the cost of any Hg emissions policy will depend on the limits (if any) placed on SO_2, NO_x, and CO_2 emissions. In most cases, tighter caps on SO_2, NO_x, and CO_2 will lower the incremental cost of satisfying a particular cap on Hg emissions. With careful alignment of the non-Hg regulations in the "Base" and "Hg Control" scenarios, EPMM allows us to isolate the incremental costs and effects of the Hg regulations specifically.

2.2 Hg Speciation

EPMM projects total Hg emissions for each year, but the fate and transport component of the framework requires that emissions be estimated in terms of the individual species, elemental Hg (Hg^0), ionic Hg (Hg^{2+}), and particulate Hg (Hg_p)[4] Thus, additional calculations are made in order to separate EPMM's total emissions projections into the individual species. These are complex calculations because the proportion of each of these species in the emitted Hg varies as a function of the type of coal being burned and the type of control equipment in place on the plant (and other non-modeled variables).

As noted above, emissions controls for SO_2 and NO_x (i.e., that were not designed to remove Hg) may have significant effects on Hg emissions. This is also true of particulate matter (PM) controls, which are employed on almost every unit in the country, and remove substantial amounts of Hg. Coal units currently incorporate one of four types of particulate emissions control equipment: out of a total of 305 GW of existing capacity, 22 GW have fabric filters (FFs), 219 GW have cold-side electrostatic precipitators (CSESPs), 48 GW have hot-side electrostatic precipitators (HSESPs), and 10 GW have Venturi-type collectors. Each of these removes most of the Hg_p in the flue gas. This explains the very small direct emissions of Hg as Hg_p from any type of generating unit. These devices are generally inefficient in removing gaseous Hg^0 or Hg^{2+}. Therefore, the ability of the Hg^0 or Hg^{2+} exiting from the boiler to convert to Hg_p by the time it reaches the PM control device affects the total Hg emissions reduction achieved by any of these PM control devices. Generally, the farther the PM control device is from the boiler (and hence the greater the residence time and cooling of flue gases before they enter the device), the greater the device's average removal of Hg before it can be emitted.[5]

---

[3] These co-controls from SCR are highly uncertain, as ICR data alone are insufficient to formulate a reasonable estimate of SCR's extent or magnitude of Hg control. Unless otherwise stated, the standard EPMM assumptions on this issue reflect the "industry consensus" developed by a working group of EPRI and utility staff engaged in Hg controls research (UARG, 2002).

[4] The ionic and particulate species deposit to the ground at much more rapid rates than elemental Hg, and thus the fate and transport step of the framework has a different relationship for each species.

[5] FFs have the highest percentage removals because they enhance the ability of the ionic and elemental Hg that enter the device to adsorb onto particles caked on the filter. Thus, in a sense FFs directly promote the formation of capturable Hg_p rather than merely remove whatever Hg_p enters the PM collection device.

Many units have or are projected to install wet FGDs to remove SO_2, and/or SCRs to remove NO_x. SCRs and wet FGDs affect the speciation of the Hg that is emitted as well as the total Hg emissions.[6] Addition of wet FGDs will tend to remove solely the ionic species. Addition of SCRs will tend to increase the fraction of ionic species while reducing that of elemental species. When SCR and wet FGD are both added, or when SCR is added to a plant with a wet FGD, then the SCRs will tend to reduce elemental emissions without a concomitant increase in ionic Hg emissions, due to capture by the FGD. Thus, "co-controls" can create a complex set of possible net changes to the speciation while reducing total Hg emissions.

In contrast, the primary control option designed specifically to address Hg emissions (i.e., sorbent injection, such as activated carbon injection) will remove both ionic and elemental species in roughly similar proportions. Thus, two scenarios with identical levels of Hg emissions could have different amounts of ionic Hg emissions, if the Hg caps are satisfied with different combinations of retrofits.

The cost-effectiveness framework's additional computations to speciate the Hg emissions are based on EPRI's analyses of the average speciation observed in emitted Hg for different unit configurations, using the ICR data (EPRI, 2000). Appendix B provides more details regarding this part of the framework calculations. The inputs to this component of the framework are EPMM's projected regional Hg emissions, and combinations of coal types and control equipment installed in the various units in the region. The outputs of this component of the framework are the regional Hg^0, Hg^{2+}, and Hg_p emissions, which sum to the total Hg emissions projected originally by EPMM.

2.3 Deposition Changes

Of the three species of Hg that are emitted, Hg^{2+} is considered the most important contributor to surface deposition close to the source, because its water solubility makes it more likely to be washed out of the atmosphere rather than to enter the global Hg cycle. (Hg_p has similar properties, but is emitted from power plants in such small amounts that it has less importance here.[7]) In contrast, Hg^0 is relatively insoluble and remains in the atmosphere for a longer period. Its average atmospheric life is approximately six months to one year, so it is subject to long-range transport beyond the source region and effectively disperses globally.

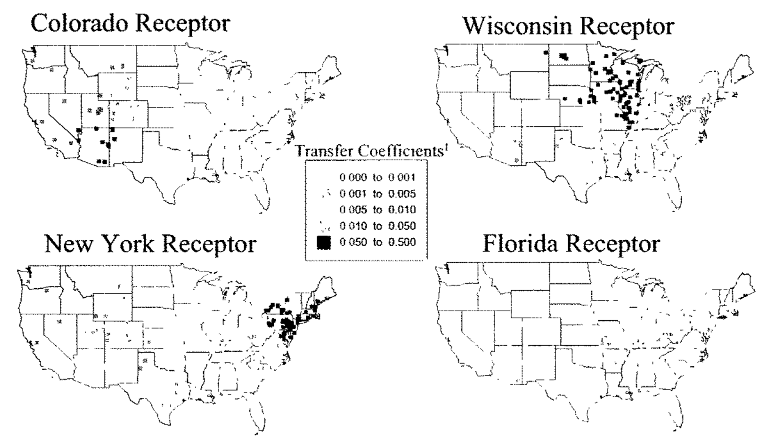
These atmospheric fate and transport relationships have been quantified by AER in its Trace Element Analysis Model (TEAM). TEAM is a three-dimensional Eulerian model that uses 1998 and 1999 data to simulate the transport, atmospheric species transformation, and wet and dry deposition of Hg over North America. The TEAM model is described in Seigneur et al. (2001), and in more detail in Pai et al. (1997). In order to incorporate the fate and transport findings of a model such as TEAM into an integrated assessment such as our cost-effectiveness framework, the model needs to be summarized into a condensed format known as a source-receptor model.

[6] Industry experts believe that the limited data suggest that selective non-catalytic reduction (SNCR) equipment, which is also designed to remove NO_x, will not significantly reduce Hg emissions.

[7] In addition, particulate-bound Hg is not believed to participate in the aquatic system-mediated conversion reaction of inorganic Hg to methylmercury. However, in this analysis we have included Hg_p in our estimation of total deposition, using the simple assumption that its emissions will vary in proportion to changes in Hg^{2+} emissions. In either case, it is such a small component of utility emissions that the specific assumption used will have no material impact on our cost-effectiveness estimates.

The key components of a source-receptor summary model are known as transfer coefficients (TCs). These constitute a set of parameters that describe how a unit of change in emissions in each of a set of specific "source regions" affects deposition in each of a set of specific "receptor locations." The TCs summarize atmospheric transformation processes as well as general transport. That is, the emissions may be defined in units of a precursor chemical, and the predicted deposition may be defined in units of a chemically transformed species of more direct environmental relevance.

AER's source-receptor representation of the TEAM model was developed to address the specific effect of the Hg emissions of coal-fired electricity generating units in the context of a global Hg cycle (Seigneur et al., 2002). It provides a separate set of TCs for emissions of Hg^0, Hg^{2+}, and Hg_p from 15 source regions of the United States (organized roughly into the boundaries of the regional electric reliability councils) to both wet and dry deposition of Hg in 19 specific receptor locations across the United States. (The 19 receptor locations are based on existing monitoring sites and a few other locations of interest to Hg researchers. Details of these receptors sites are provided in Appendix C.) There is also a TC for total U.S. deposition changes due to emissions from any one source region. Figure 2-2 presents a graphical summary of the TCs associated with four of the 19 receptor locations.



[1]Total Deposition / Total Hg Emissions (µg/m²-Mg)

Figure 2-2
Transfer Coefficients for Four Receptor Locations

(TCs are stated in units of total deposition at the designated receptor (in µg/m²) per Mg of Hg emitted from each source shown as a dot on the map. Color of each dot reflects the magnitude of the impact of Hg emissions from each source on the receptor in question. Blue indicates the least source-receptor impact, and red indicates the highest source-receptor impact.)

These TCs allow our cost-effectiveness framework to reflect the atmospheric settlement and gradual deposition of Hg emissions that are projected in the preceding two steps. The TCs used in the framework are stated as a percentage change in regional deposition per percentage change

in 1998/1999 emissions from each source region. Each region's future projected annual emissions of Hg^0 and Hg^{2+} (i.e., the outputs of the previous framework step, derived from the EPMM model run) are converted into a proportion of 1998/1999 emissions for the same region, then multiplied by the appropriate TC.[8] For the current version of our framework, we do not differentiate between wet and dry deposition. Rather, we sum wet and dry deposition to obtain the total Hg deposition change in each receptor location.

According to AER's analysis of 1998/1999 data, electricity generating units were responsible for less than 8% of Hg deposited in the United States, so the effect of reductions in electricity generation Hg emissions on U.S. Hg deposition is expected to be modest. For example, a 10% reduction in national ionic Hg emissions from utility sources results in only a 0.75% reduction in U.S. Hg deposition, and a 10% reduction in national elemental Hg emissions lowers U.S. deposition by only 0.03%. Thus, a relatively large percent reduction in national Hg emissions from electricity generation may have fairly little impact on overall Hg deposition levels. However, the changes can vary more significantly from receptor to receptor, and this variability is retained in the framework's calculations.

We do not have data at present on the geographical source distribution of the wild freshwater fish that are consumed in the United States. Given this uncertainty, it is difficult to construct a single overall metric of deposition change that would be directly proportional to change in MeHg exposure via consumption of wild freshwater fish. Thus, the framework reports the average of the three largest and three smallest regional deposition changes as well as a simple average over all regions.[9]

AER has not developed TCs for deposition to the world's oceans. U.S. electricity generating emissions are a much smaller proportion of deposition in these more distant locations than in the continental United States – they account for only about 46 of 6,000 tons (0.8%) of Hg currently estimated to be deposited globally each year. Nevertheless, we account for these small deposition changes globally because ocean fish are the largest source of MeHg in the average U.S. individual's diet. For this portion of deposition changes, we simply assume that all ocean fish respond equally to a given percentage reduction in U.S. electricity generating emissions. There is tremendous uncertainty regarding which oceans and which types of ocean fish consumed in the United States will most respond to U.S. Hg emissions reductions. Our simple assumption at least avoids neglecting this potential additional reduction in exposure to U.S. citizens from domestic control steps. By including it, we are also able to understand better where the largest uncertainties are in estimating exposure and risk reductions from lowered U.S. Hg emissions.

---

[8] Only around 1% of electricity generating unit Hg emissions is in particulate form, so we make the simplifying assumption that the percentage change in Hg_p emissions is equal to the percentage change estimated for Hg^{2+} emissions.

[9] The simple average change in deposition is used as the metric for assessing changes in MeHg exposure from consumption of domestic wild freshwater fish on the population. The consumption rate of wild freshwater fish is estimated to be equal across each of the modeled receptor locations. With better information on the share of all wild freshwater fish caught and consumed in each geographical area and with data on the distribution of the consumption of wild freshwater fish across the population, we could construct a weighted average that might better reflect the distributional effect on exposures. In the absence of such data, it is useful to consider the range of deposition changes across all the regions as well.

2.4 Changes in Human Mercury Exposure

Once we have estimated Hg deposition changes in various regions, the next step in our cost-effectiveness framework is to estimate the change in methylmercury (MeHg) exposure to the population of women of childbearing age. Of particular interest is the change in exposure to women at the upper end of the MeHg exposure distribution, that is women who are exposed to MeHg at levels above EPA's RfD. The following calculations enable us to estimate the effect reductions in mercury emissions from U.S. coal-fired power plants have upon the exposures of U.S. women to MeHg, including the specific effect of Hg controls in reducing the number of women with exposures above the RfD.

To estimate the effect of a change in Hg deposition on the MeHg consumption of the at-risk population, we make several assumptions. First, we assume that all MeHg exposure occurs through fish and shellfish consumption. This assumption is consistent with the conclusions in EPA's *Mercury Study Report to Congress* (EPA, 1997). The second assumption is that the change in MeHg content of all fish and shellfish in a given year is proportional to any change in Hg deposition in that year. That is, the assumption is made that any reductions in Hg deposition will immediately be reflected by comparable reductions in MeHg concentration in fish. This is a high-end assumption; in fact, the MeHg content of fish will decline only gradually if Hg deposition is reduced, because the concentration of Hg in lake and river sediments will depend on Hg deposition in prior years, and the sediment concentrations will not instantaneously decline. In addition, the MeHg concentration of fish, and particularly the high MeHg concentration of top predator fish, will decline only as entire food chains re-equilibrate to reflect the reduction of mercury inputs to each aquatic ecosystem.

In the analysis below, three separate sources of MeHg consumption are considered. First are wild, freshwater fish. These fish are assumed to experience reductions in MeHg in proportion to the change in deposition averaged over the lower 48 states. As described, deposition rates are sensitive to the mix of elemental versus ionic mercury emissions.

Second, it is assumed that the MeHg content of marine fish will decline in proportion to changes in total global mercury emissions brought about by changes in U.S. utility emissions of mercury alone. For marine fish, the speciation of Hg emissions is not considered; it is assumed that the concentration of MeHg in marine fish is proportional to total mercury emissions, including both anthropogenic and natural sources. For the purpose of these calculations, an annual global emission rate for total mercury of 6,000 tons per year from anthropogenic and background sources combined will be assumed; this was the value in AER's deposition calculations.

Third, the assumption that farm-raised fish do not contribute to MeHg exposures was made. While this assumption may not be completely accurate, it is roughly correct for catfish and trout, the largest farm-raised fish crops (Santerre et al., 2001). The assumption that farm-raised fish do not contribute to MeHg exposure is conservative in the sense that all MeHg intake is attributed to either wild freshwater fish or to marine fish, with concentrations in those fish assumed to link instantaneously to deposition.

Fourth, we have assumed that MeHg exposures from the consumption of marine and wild, freshwater fish are proportional to the amount consumed. This is a simplifying assumption because, for a given amount of MeHg in an aquatic ecosystem, MeHg concentrations will tend to be higher in the larger fish that are higher on the food chain in that ecosystem; additionally, the response of those concentrations will also likely differ in response to a given change in MeHg in the water. This assumption means that the reduction in MeHg exposure associated with reduced emissions and deposition will be proportional to the reduction in MeHg in marine fish and in freshwater fish, each weighted by the relative fraction of each type consumed. That is, if 90% of consumption of fish with significant levels of MeHg is of marine fish and 10% is of wild freshwater fish, then we assume that 90% of MeHg exposure comes from marine fish and 10% from wild, freshwater fish. Because it is assumed that farm-raised fish do not contribute to exposure, the fraction of fish consumed that are farm-raised does not affect the calculation.

The approach described above assumes that the fish consumption patterns of women of childbearing age are not affected by changes in mercury emissions. It is possible that reductions in mercury emissions and deposition could lead to fewer water bodies with fish consumption advisories based on MeHg. If this were to occur, then women might eat more wild fish, and reduction in MeHg concentrations in wild fish might be partially offset by increased consumption.[10] Such a possibility is not considered in this analysis.

The available data concerning fish consumption, particularly consumption of wild freshwater fish, are not adequate to support an analysis that takes several distributional patterns into consideration. For example, the relative fraction of marine and freshwater fish consumed by women in the upper end of the fish consumption distribution varies from one woman to the next. For those women who consume a greater than average percentage of fish that are wild freshwater fish, the effect of reductions in mercury emissions will be greater than for women who eat only marine fish. However, it is difficult to get a clear picture of freshwater fish consumption by women in the survey because the total number of meals reported for "other" or "unknown" fish exceeds the number of meals reported for freshwater fish. Similarly, estimates of changes in MeHg consumption are not made on a regional basis because geographic distribution of consumption of wild freshwater fish is not refined.[11] In addition, data on the geographic distribution of blood levels of mercury are not available, so we cannot determine whether women with high blood concentrations live in areas where above-average quantities of wild, freshwater fish are consumed.

Given the assumptions described here, it is necessary to estimate the relative consumption rates for marine and wild fish. Data concerning fish consumption by women between the ages of 16 and 49 has recently been released by the U.S. Centers for Disease Control and Prevention (2002). These data were developed as part of the *National Health and Nutrition Examination Survey*

[10] Such feedback effects are further complicated by the fact that natural water body carrying capacities are driven by productivity parameters; increased fishing due to improvements in Hg burdens may have an impact on the residual fish populations.

[11] Also, sensitive aquatic resources, such as low pH lakes, wetlands, salt marshes, the Everglades, etc., are not randomly distributed in North America. As a result, changes in Hg burdens in fish from these waters are unlikely to be accurately simulated by a regional impacts analysis such as this, without much greater detail in the data than are currently available to this analysis.

(NHANES) conducted by the National Center for Health Statistics. The recent NHANES data release includes blood Hg concentration data on 1,709 women, of whom 286 were pregnant at the time of the survey and blood sampling. Questionnaires regarding fish consumption were also completed as a part of NHANES. 1,646 questionnaires are available for women with blood Hg data, including 266 of the pregnant women. Details on the NHANES fish and blood data are provided in Appendix D, and are briefly summarized below.

The NHANES questionnaires asked women how many fish meals, split into twenty-one categories, they had consumed over a thirty day period. However, several categories – breaded fish, other fish, and unknown fish – cannot be identified as either marine or freshwater fish. Combined, these three categories make up slightly over 25% of all responses in the survey. If "other" and "unknown" are excluded, 384 fish meals were reported for freshwater fish, out of 3,541 total reported meals. Of these 384 meals, 302 were catfish or trout, the two fish species that are most commonly farm-raised. In addition, 341 meals of salmon were reported. Salmon, an estuarine fish, is also increasingly farm-raised. Overall, these data do not unambiguously identify the relative fractions of marine and wild freshwater fish.

Fish consumption data from a variety of data sources are reported in the EPA's *Mercury Study Report to Congress*, in a 2002 EPA report, and in EPA's *Exposure Factors Handbook* (U.S. EPA, 1997b, 2002, and 1997a, respectively). Table 10-7 of the *Exposure Factors Handbook* indicates that the mean consumption rate of finfish (that is, not including shellfish) from freshwater and estuarine habitats is 3.6 grams per day, and from marine habitats 12.5 grams per day. However, the freshwater and estuarine habitat estimate includes farm-raised catfish, trout, and salmon, which most likely accounts for more than half of the total. The U.S. Department of Agriculture (2002) estimates that 2002 sales of domestic farm-raised catfish to processors would be in the range of 603-615 million pounds, with sales by processors of fillets and other cleaned products about half this amount. In addition, this document estimates that 360 million pounds of Atlantic salmon and 125 million pounds of tilapia will be imported in 2002, of which a large portion represents fillets rather than whole fish. USDA also indicates that a significant but unspecified amount of catfish was imported from Vietnam. USDA estimates that $64.5 million of food-size trout will be sold in 2002, but did not translate this into units of weight. In any case, from these estimates it appears that roughly a billion pounds of farm-raised fish were sold in the United States. If one roughly assumes that half this amount was consumed (allowing for the inedible portion of whole fish), this translates into about 2.2 grams per day per capita. On this basis, the assumption that wild, freshwater fish make up 10% of fish consumed when farm-raised fish are excluded appears reasonable.

Given the data described above, we assume that 90% of the contribution to MeHg exposure comes from marine fish consumption, and the remaining 10% is due to consumption of wild, freshwater fish. To calculate the percent change in exposure resulting from a percent change in Hg emissions and deposition, we use the following formula:

$$\Delta(\text{exposure}) = 0.1 * \{\Delta(\text{deposition})[\textit{as a fraction}]\} + 0.9 * \left\{\frac{\Delta(\text{emission})[\textit{in tpy}]}{6{,}000\,\textit{tpy}}\right\} \qquad \textbf{Eq. 2-1}$$

2.5 Characteristics of Methylmercury Exposures and the Effect of Exposure Reductions on the Number at Risk

Exposures to MeHg are regulated by the EPA based on the reference dose (RfD), which the EPA defines as "an estimate (with uncertainty spanning perhaps an order of magnitude) of a daily oral exposure to the human population (including sensitive subgroups) that is likely to be without an appreciable risk of deleterious effects during a lifetime. It can be derived from a No Observed Adverse Effect Level (NOAEL), Lowest Observed Adverse Effect Level (LOAEL), or benchmark dose (BMD), with uncertainty factors generally applied to reflect limitations of the data used." The EPA RfD for MeHg is based on the benchmark dose limits calculated for a number of developmental tests, primarily but not exclusively from the Faeroes Islands study. The benchmark dose's lower confidence limit was divided by an uncertainty factor of 10 to set the RfD. The resulting RfD for MeHg is 0.1 μg/kg-day, which defines an intake rate in units of micrograms of MeHg per day per kilogram of body weight. In the documentation supporting this RfD,[12] analyses provided indicate that an intake rate at the RfD is calculated to produce a blood concentration of MeHg of 5.37 μg/L, equivalent to 5.37 ppb.

As described in Appendix D, the current levels of exposure of U.S. women to MeHg can be determined from the recently released NHANES data. The NHANES data indicate that about 8% of U.S. women between the ages of 16 and 49 are exposed above the RfD, but that the degree by which the RfD is exceeded is small. None of the 1709 women tested in the NHANES study had blood levels above the benchmark response level on which the RfD was based; only about 2.5% had blood levels twice those of the RfD, and the highest MeHg level measured was slightly below five times the RfD. Given that all U.S. women in the NHANES dataset had blood levels within the range of the RfD or below, and that the RfD level is defined to be without appreciable risk, the risk to U.S. women from MeHg, given the low exposure levels that apparently exist, is by no means certain.

Effects observed in the Faroe Islands study, on which the RfD is based, were to a population exposed to MeHg at levels approaching an order of magnitude higher than those of U.S. women. The exposure levels of U.S. women is significant lower than that of populations where adverse effects from MeHg exposure have been reported. The mean cord blood concentration in the Faroes study was 20.4 μg/L. When adjusted for the observed result that cord blood is, on average, 1.7 times that of maternal blood, this indicates a mean blood MeHg concentration in Faroese women of 12 μg/L. In contrast, the mean concentration of total mercury (which is higher than the concentration of MeHg) in the NHANES data is 1.7 μg/L. [Too many measurements of speciated mercury in the NHANES study were below the detection limit to allow calculation of mean MeHg exposure levels.] For comparison, the mean hair concentration in the Seychelles study, in which no adverse response was observed, corresponds to a blood concentration of MeHg of 32 μg/L.

To quantify the effectiveness of exposure reductions, we use an exploratory measure: the reduction in the number of babies born annually in the United States to mothers with blood MeHg levels above the RfD. This is estimated by assuming a percentage shift in the exposure of each individual fish consumer that is equal to the average percentage change in estimated MeHg

[12] Documentation is from EPA's Integrated Risk Information System (IRIS), at http://www.epa.gov/iris/subst/0073.htm.

consumption. We then estimate the impact of these changes on the distribution of MeHg blood levels for women of child-bearing age, focusing on the fraction of the relevant population above the RfD. In this way, we can estimate the number of women (and the associated expected number of pregnancies) with exposures above the RfD under an emissions reduction scenario and under the base case. No attempt is made to quantify or monetize the health benefits as was done previously in the analysis of lead exposure and its effect on IQ.

Information supporting these estimates is developed from the NHANES data, using its reported distributions from sampling of blood MeHg levels in women of childbearing age. NHANES provides blood MeHg concentration data for women of childbearing age. When appropriate weights to adjust for demographic factors are applied (that is, when the data are weighted to make the observations representative of U.S. women in terms of age, race, and geographic location) the cumulative distribution of the blood MeHg levels can be estimated. In order to calculate the effect of a percent change in blood MeHg levels on the distribution relative to the RfD, we calculated the slope of the distribution in the region near the RfD (that is, within 0.5 ppb of the RfD value of 5.37 ppb), as shown in Figure 2-3. This regression yields the following equation:

$$y = 39.527x - 31.107, \qquad \textbf{Eq. 2-2}$$

in which x is the percentile of the distribution and y is the blood MeHg level in parts per billion. This linear fit to the data in the region of the RfD indicates that the RfD blood level corresponds to the 92.3 percentile of the distribution. This means that, for the population sampled in the NHANES study, weighted to adjust for demographic factors, 7.7% of women aged 16-49 have blood levels above the RfD.

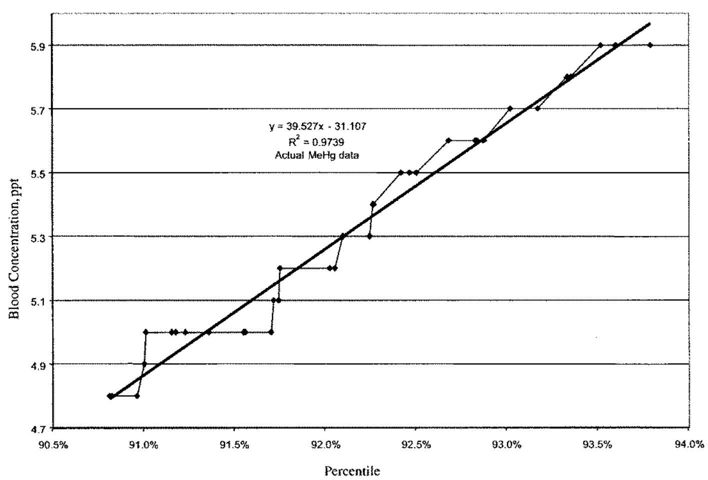


Figure 2-3
Regression on NHANES Data to Estimate Slope of Distribution Near RfD

Additionally, if we assume that blood MeHg levels will respond in direct proportion to changes in MeHg consumption, then we can use the above relationship to estimate the change in the number of women that have blood MeHg levels above the RfD if mercury emissions are reduced. That is, if we decrease the blood MeHg levels of each member of this population by 1%, the change in the percentage of women of childbearing age whose blood MeHg levels exceed the RfD can be calculated based on the slope of the distribution:

$$[(1\%)*(RfD)]/(39.527) = 0.136\% \qquad \textbf{Eq. 2-3}$$

This change can then be converted into an estimate of the change in the number of babies born each year to women with blood MeHg levels above the RfD by multiplying the change in the percentage of women above the RfD by the projected number of annual births in the United States.

In our current framework calculations, we use projections of annual births in the United States from the Census Bureau's 2000 "Highest Series,"[13] which assumes high rates of population growth due to high fertility rates, high rates of net immigration, and rapid increases in life expectancy. As a point of reference, there are roughly 4 million live births per year in the United States at present. The Census "High Series" projects this will increase to 5.7 million births per year by 2020.

2.6 Cost-Effectiveness: The Final Step

The estimated change in the number of births per year to women with blood MeHg concentrations above the RfD can be compared to the estimated change in the cost of electricity production due to mercury controls alone, in order to calculate the average cost of reducing the number of such "at risk" births. One could also consider cost-effectiveness metrics such as dollars per percent change in deposition, or per percent change in total consumption.

Another way of comparing among scenarios is to consider how two alternative policies differ in terms of where the deposition changes occur. If there are no significant differences in amount of change in each region, then the less costly policy should be preferable. In a situation where legislative changes are being discussed, this might be useful in comparing the implications of allowing emissions trading as a way of reducing the costs of achieving Hg reductions. It is also useful for considering the merits of allowing "subcategorization" in a MACT standard, in a situation where only Section 112 implementation is under consideration.

Note that the cost-effectiveness framework does not attempt to monetize the ecological changes or the reductions in human health risk; it is not a *cost-benefit* framework. Adding monetization would be an easy step to implement computationally, but extremely limited in its usefulness, given the current lack of any technically justifiable inputs for such a computation.

In Section 3, we apply our framework to a pair of publicly proposed utility mercury control proposals. These proposals were each advanced in mid-2002, in the U.S. Congress and in a public working group convened by the U.S. EPA, respectively. The purpose of these calculations is to illustrate the steps in the framework, and to provide a preliminary assessment of several cost-effectiveness measures for the two alternative approaches to utility Hg controls.

[13] Data found at http://www.census.gov/population/www/projections/popproj.html

3
ASSESSMENT OF TWO ALTERNATIVE MERCURY CONTROL POLICIES

Two basic types of Hg control policies have recently been proposed. One is a Hg cap-and-trade program, implemented as part of new multi-pollutant legislation. The other is to regulate Hg under Section 112 of the Clean Air Act with a Maximum Achievable Control Technology (MACT) standard. In our illustrative application of the cost-effectiveness framework, we will address a single example of each:

1. The Hg controls of the Clear Skies Act of 2002 (the "CSA" scenario).
2. A MACT standard set at 2.2 lbs/10^{12}Btu (trillion Btu, or "tBtu") stack limit (the "MACT" scenario).[14]

In order to maintain comparability, we need a single "Base" scenario that reflects co-existing policies for the non-Hg emissions of electricity generation. For the Base scenario of the following illustrative analysis, we will use the SO_2 and NO_x caps that are also expressed in the Clear Skies Act (CSA).[15] This Base scenario allows us to isolate the incremental impacts and costs of the Hg control provisions of CSA. Additionally, since it contains caps on SO_2 and NO_x that may occur even if multi-pollutant legislation is not passed (due to implementation of the $PM_{2.5}$ air quality standard), it allows us to assess the incremental costs and effects of a MACT standard in the context of a likely set of future SO_2 and NO_x controls.

Accounting for the co-existing SO_2 and NO_x limits is crucial because of the potential impact of Hg "co-control" from retrofits that reduce SO_2 and NO_x emissions. The following analysis focuses on the *incremental* reductions in risk and increases in cost engendered by the addition of explicit Hg control requirements on top of the Base scenario's co-controls:

- The CSA scenario implements a national cap-and-trade program with total Hg emissions capped at 26 tons in 2010, then at 15 tons in 2018. Note that the phasing of the CSA Hg cap is in general alignment with the phasing of the Base scenario's caps on SO_2 and NO_x.
- The MACT scenario imposes its Hg requirements fully by 2008, consistent with that specific regulatory process's current schedule. The example of a MACT used here is a 2.2 lb/tBtu emissions rate limit applied at every unit (without subcategorization). No trading or averaging across plants is allowed in this scenario.

---

[14] "tBtu" = trillion Btu = 10^{12} Btu.

[15] For SO_2, the existing 8.95 million ton per year cap is reduced to 4.5 million tons in 2010, and to 3 million tons in 2018. For NO_x, an annual cap is applied to the East and West, with national emissions being reduced to 2.1 million tons in 2008, then to 1.7 million tons in 2018.

This illustrative MACT scenario could be viewed as "less stringent" than the CSA cap, because if each unit were to exactly meet this MACT limit, total emissions would likely be about 24 tons per year, rather than 15 tons as in the CSA. On the other hand, the MACT is fully implemented in 2008, well ahead of any co-controls that might be obtained from the need to reduce SO_2 and NO_x emissions as well.

3.1 Clear Skies Act

We begin our illustrative examples of framework calculations by running EPMM with the CSA emissions controls, and with the "Base" emissions controls (i.e., CSA minus its Hg controls). Table 3-1 displays the changes in annual coal consumption for each of 10 generic coal types projected by EPMM for the CSA compared to the Base scenario. Total coal consumption falls increasingly as the cap tightens, with 402 tBtu less coal consumption (on a heat content basis) by 2020 (this is about 2% less than in the Base scenario). The Hg controls in the CSA scenario result in some substitution out of coal-fired generation (primarily into gas-fired generation), which accounts for the decrease in coal consumption.

Table 3-1
Differences in Coal Consumption in Trillions of Btu (CSA Compared to Base)

b	Bituminous							Subbituminous		Lignite	Total
Sulfur Content	Low		Med		High			Low	Med		
Hg Content	Low	Med	Low	Med	Low	Med	High	Low			
2004	0	0	0	0	0	54	-24	0	0	0	30
2008	0	0	177	-36	0	-115	-296	0	0	0	-270
2010	0	0	552	-298	0	95	-474	0	0	-22	-148
2012	0	0	405	-279	0	146	-303	0	-18	-24	-73
2015	0	0	646	-331	313	-91	-401	0	-207	-54	-125
2018	315	0	850	-374	571	-85	-513	-822	-76	-23	-157
2020	0	0	36	-37	1,427	-52	-1,469	-286	-22	0	-402

The CSA Hg caps also induce a modest amount of substitution among coal types. Table 3-1 shows that, in 2020, the Hg caps in the CSA induce about 1,469 tBtu less consumption of high-sulfur/high-Hg bituminous coal, which has the highest average Hg content of the ten generic coal types (EPMM assumes this type of coal contains 16.2 lbs Hg/tBtu). This is replaced primarily by other high-sulfur bituminous coals, although with a much lower average Hg content (e.g., 7.3 lbs/tBtu). Thus, switching appears to be primarily from one type of bituminous coal to another that has relatively a much lower Hg content. It appears that most of the net reduction in total coal consumption is borne by subbituminous coal, which has a lower average Hg content, but also a relatively low average level of co-control.

Retrofits play the largest role in satisfying the CSA limits on Hg emissions. Tables 3-2 and 3-3 show the amount of each type of retrofit installed in each year for the two scenarios. There are two possible types of direct Hg control technologies in the model. Both are based on activated carbon injection (ACI), but one is operated at a carbon injection rate designed to remove only 60% of the Hg while the other operates at a 90% removal rate. Because there are no Hg limits in the Base scenario, and because ACI does not remove NO_x or SO_2, none of the Hg control devices is installed in the Base scenario (Table 3-2). In the CSA scenario (Table 3-3), over 122 GW of capacity (about 40% of the existing coal capacity) installs some form of ACI system by 2020.[16]

The amount of capacity that adds FGDs and SCRs also changes as a result of the Hg controls in the CSA scenario. Because coal consumption, particularly the consumption of high-sulfur coal, is lower in the CSA scenario, fewer FGDs are required to satisfy the SO_2 limits. There are also fewer SCRs installed, because the reduction in coal consumption reduces the need for equipment that removes NO_x. However, all of the avoided FGDs and SCRs are those that would have been installed after 2010. Many of the earlier retrofits still make sense because they help meet the first phase of the Hg cap, while also providing bankable excess SO_2 and NO_x allowances. Thus, one can see how co-control can create interactions in compliance strategies across multiple pollutants. Recall that the types of controls applied to reduce the emissions will affect the *species* of Hg that are reduced, as we will explore after next discussing the CSA control cost results.

Table 3-2
Capacity (GW) Retrofitted in Base Scenario

Year	FGD	SCR	60% Hg Control	90% Hg Control
In place as of 2000	87,612	7,891	0	0
2004	8,562	81,198		
2008	27,458	68,834		
2010	12,126	7,048		
2012	2,333	6,168		
2015	11,451	12,216		
2018	24,118	14,351		
2020	48,103	9,310		
Total Added	134,151	199,126	0	0

[16] All of the Hg controls in the CSA case are selected at the 90% control level. This is because the cap-and-trade feature of the CSA Hg caps induces controls with the lowest $/ton removed cost. The 90% control level is always lower in terms of $/ton removed than the 60% control level on any given unit.

Table 3-3
Capacity (GW) Retrofitted in CSA Scenario

Year	FGD	SCR	60% Hg Control	90% Hg Control
In place as of 2000	87,612	7,891	0	0
2004	6,927	81,198		
2008	27,450	65,236		13,479
2010	16,630	10,668		56,093
2012	2,227	4,900		4,651
2015	15,050	10,215		13,882
2018	33,979	18,183		17,909
2020	20,346	4,371		16,336
Total Added	122,609	194,770	0	122,350

Table 3-4 shows the incremental electricity generation cost of the CSA compared to the Base Case. This cost has four main components: the increased capital cost of installing retrofits; the increased cost of operating retrofits (the change in Variable O&M Costs); the increased cost of burning gas and higher-cost coal rather than lower-cost coal (the increase in Fuel Costs); and the increased cost of building combined cycle plants rather than combustion turbine plants to satisfy reserve margins (New Plant Capital Charges and New Plant Fixed Costs).[17] After 2010, which is when the Hg cap is first imposed, costs are more than $1 billion per year higher in the CSA scenario than in the Base scenario. These are the incremental costs of *only* the Hg provisions of the CSA, because the Base scenario already contains the CSA's NO_x and SO_2 caps and therefore also the costs associated with those non-Hg provisions of the CSA.

The above annual incremental costs are one key element of a cost-effectiveness analysis. Next we discuss the effectiveness of the CSA's Hg controls: changes in emissions, deposition, MeHg exposure, and human risk.

The CSA has lower total emissions than the Base case, but it is important to note that even our Base scenario has a significant downward trend in emissions. Figure 3-1 shows total Hg^0 and Hg^{2+} emissions in each modeled year for the Base scenario, and Figure 3-2 shows these emissions for the CSA scenario.[18] Base scenario emissions (Figure 3-1) are about 46 tons in 2004, but fall to about 37 tons by 2020, without a dedicated Hg constraint, and despite a 33% increase in total generation. This is a direct result of the co-controls from meeting the SO_2 and NO_x caps. Notice also that all of the Base scenario emissions reduction is in the Hg^{2+} species, which is the more important species from the perspective of affecting domestic deposition.

[17] Combined cycle capacity is more expensive but more efficient than combustion turbine capacity. Relatively more of the former type capacity is built in the CSA scenario, because it can displace coal generation, which reduces Hg emissions.

[18] These U.S. total emissions are actually calculated for each of the 28 EPMM market regions.

Table 3-4
CSA's Incremental Cost Relative to Base Case ($millions in 1999 Dollars)

	2008	2010	2012	2015	2018	2020
Annual Capital Charges						
New Plant	$88	-$34	$13	$49	$49	$118
Retrofits	-$36	$484	$503	$669	$1,164	$402
Sub-Total	$51	$451	$516	$718	$1,213	$521
Fixed Operating Costs						
Existing Plant	-$11	$5	-$1	-$5	-$6	-$6
New Plant	$18	-$2	$6	$11	$11	$35
Retrofits	-$7	$82	$84	$125	$240	$15
Sub-Total	$0	$85	$89	$132	$245	$44
Variable Operating Costs						
Variable O&M Cost	$28	$350	$365	$442	$622	$463
Fuel Cost	$199	$243	$113	$158	$145	$397
Sub-Total	$227	$594	$478	$600	$767	$860
Total Incremental Cost	$278	$1,129	$1,083	$1,450	$2,225	$1,425

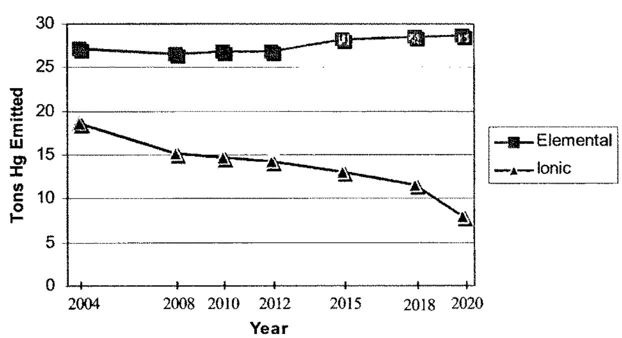


Figure 3-1
Hg^0 and Hg^{2+} Emissions Projections – Base Case

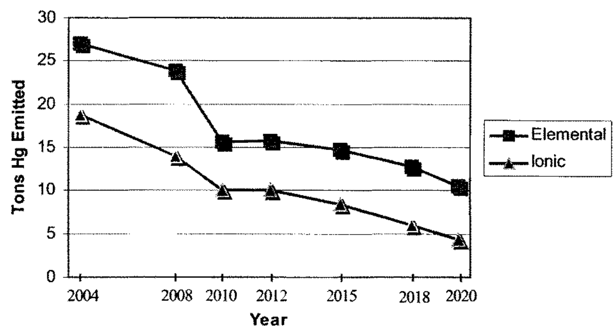


Figure 3-2
Hg^0 and Hg^{2+} Emissions Projections – CSA Case

In the CSA scenario (Figure 3-2), total Hg emissions are 22 tons lower in 2020 than in the Base scenario (15 tons vs. 37 tons). However, most of this reduction comes from the Hg^0 species, and the difference in ionic Hg emissions is fairly small. The large difference between the speciation of Hg reductions caused by CSA's Hg controls and the speciation of the reductions caused by the SO_2 and NO_x controls is due to the difference in the means of compliance. In the Base scenario, FGDs are retrofitted onto 134 GW (almost 45%) of coal capacity. Because these devices primarily remove ionic Hg, most of the Hg emissions reductions that result are in the ionic form. When the incremental requirements of Hg caps are imposed in the CSA scenario, they are satisfied primarily by Hg-specific retrofits (ACI) and, to lesser degrees, by switching from high-Hg coal to lower-Hg coal and by switching to relatively more use of gas. All of the latter control measures reduce elemental and ionic Hg emissions in equal proportions. Thus, given that most of the Hg emissions remaining after the installation of Base scenario retrofits are Hg^0, only a small incremental reduction of Hg^{2+} emissions results from the imposition of an explicit Hg cap in the CSA.

The next step is to convert these emissions changes into deposition changes. For each year, we take the projected elemental and ionic Hg emissions in each of the 15 emissions source regions and, using the TCs derived from the TEAM model, calculate the difference in Hg deposition at the 19 receptor sites. In 2020, for example, total elemental Hg emissions are about 18 tons lower in the CSA scenario than in the Base scenario, and ionic emissions are about 4 tons lower. This results in an average 1.5% decrease in total U.S. Hg deposition.[19] Table 3-5 provides the resulting percent differences in deposition of the CSA compared to the Base scenario in each of the 19 receptor locations. The change ranges from -0.1% to -5.3% across the receptor sites in 2020.

[19] This is a simple average over the changes estimated for each of the 19 receptor sites, and is used only for illustrative purposes. Actual risk changes need to account for variation in changes across the continent, while also weighting for the relative importance of deposition in each area. Until appropriate weighting factors are available, we feel it best to use simple averages to summarize overall differences between scenarios.

Much of the attention regarding risks of human exposure to MeHg is focused on consumption of contaminated wild fish from the lakes and streams of the United States. However, as noted in Section 2.4, wild freshwater fish account for a small fraction of total U.S. fish consumption. Some believe that the reduction of Hg in freshwater sources is critical to protect populations where fishing accounts for a large portion of their diet. However, such populations are both very small and not well documented. The more typical pattern of MeHg exposure in the U.S. population would seem to be primarily marine fish (90% of the relevant consumption), with wild freshwater fish a distant second (10% of the relevant consumption). Our analysis of exposure reduction through Hg control policies therefore considers both freshwater and marine fish sources.

Table 3-5
Domestic Deposition Changes by Receptor Site (CSA compared to Base)

	2008	2010	2012	2015	2018	2020
AL00	-0.5%	-3.6%	-3.4%	-5.2%	-5.0%	-0.5%
CO00	-0.4%	-1.9%	-0.3%	-1.4%	-1.4%	-0.6%
FL00	-0.1%	-0.3%	-0.3%	-0.4%	-0.5%	-0.2%
FL01	-0.7%	-0.4%	-0.5%	-1.0%	-1.0%	-0.6%
FL11	0.0%	-0.1%	-0.1%	-0.1%	-0.1%	-0.1%
GA00	-0.2%	-1.2%	-1.2%	-1.7%	-1.8%	-0.7%
GM00	0.0%	-0.1%	-0.1%	-0.1%	-0.1%	-0.1%
LA00	-0.2%	-1.6%	-1.5%	-2.3%	-2.2%	-0.3%
ME09	-0.1%	-0.5%	-0.4%	-0.5%	-0.6%	-0.6%
MI00	-0.6%	-2.4%	-2.1%	-2.4%	-2.6%	-1.9%
NH05	-0.8%	-3.1%	-0.7%	-2.3%	-2.4%	-1.0%
NJ00	-0.3%	-1.0%	-1.2%	-1.1%	-1.3%	-1.2%
NY20	-0.3%	-1.3%	-1.1%	-1.2%	-1.5%	-1.3%
PA30	-0.7%	-3.8%	-3.5%	-3.4%	-4.2%	-3.5%
TN00	1.0%	-1.7%	-1.5%	-2.6%	-3.7%	-3.8%
TX00	-0.1%	-0.3%	-0.3%	-0.3%	-0.3%	-0.1%
TX21	-3.9%	-8.3%	-9.6%	-8.5%	-8.8%	-5.3%
WI08	-0.4%	-1.5%	-1.3%	-1.5%	-1.6%	-1.4%
WI31	-1.5%	-5.6%	-4.9%	-5.5%	-5.5%	-4.4%
Average	-0.5%	-2.0%	-1.8%	-2.2%	-2.3%	-1.5%

The percentage changes in deposition summarized in Table 3-5 reflect the opportunities for reducing exposures to MeHg from consumption of wild freshwater fish caught in the continental United States. The actual change in freshwater fish concentrations may vary from 0 to about 5% by 2020, but for the following illustrative computations we will use the simple average changes across all of the receptor sites. Based on our assumption that wild freshwater fish account for 10% of the blood MeHg content of the at-risk population, this implies a (1.5%)*(10%) = 0.15% reduction in the at-risk population's blood MeHg content. Our analysis of the distribution of blood MeHg levels of women of childbearing age in the NHANES survey implies that a 0.15% reduction in blood levels of MeHg will reduce the fraction of those women with blood MeHg levels above the RfD from 7.68% to 7.66%. In other words, the fraction of births by mothers with potentially risky MeHg levels falls by 0.02 percentage points. The U.S. Bureau of Census projects about 5.7 million births in the United States in 2020, so the number of births moved below the RfD level from freshwater Hg reduction is 1,118. Of course, if all the wild freshwater fish consumed in the United States came from regions such as the average of the least-changed three in Table 3-5, or from the average of the most-changed three, excluding the largest variants, the incremental reduction in births by women with blood levels above the RfD due to reduced exposure to MeHg through consumption of wild freshwater fish could range from 101 to 3,023.

To the above risk reductions we add the changes in MeHg risk from consumption of marine fish. Although marine fish appears to account for about 90% of the MeHg exposure of the U.S. population, the deposition changes to the marine environment due to U.S. emissions changes will be much smaller than those in the continental United States, shown in Table 3-5. Without an oceanic deposition model, we assume marine MeHg concentrations globally will fall by a percent equal to the current U.S. utility emissions contribution to total global emissions, multiplied by the percent change in those emissions. Given the global nature of this deposition, we assume that a ton of elemental Hg is equally likely to be deposited as a ton of ionic Hg. Global deposition has been estimated at approximately 6,000 tons annually. Therefore, a 10-ton reduction in total U.S. utility emissions is assumed to produce a 10/6,000 (0.17%) reduction in deposition in oceans globally. Based on our assumption that marine fish account for 90% of the blood MeHg content of the at-risk population, this implies a (0.17%)*(90%) = 0.15% reduction in the at-risk population's blood MeHg content and, as mentioned previously in our freshwater example, a corresponding decrease in the fraction of women exposed. In 2020, the total Hg emission change in the CSA scenario is 22 tons, resulting in 2,508 incrementally protected births. Note that this is more than twice the estimated 1,118 risk reductions due to MeHg reductions in U.S. wild freshwater fish. The risk reduction from marine fish is, of course, subject to even more uncertainty than that for freshwater fish.

Table 3-6 provides a summary of national average cost and effectiveness results for each of the years modeled for the CSA scenario. The incremental costs shown represent the net change between the CSA and Base scenarios. Hg emissions and Hg deposition totals are shown for the CSA scenario, along with the percentage change in deposition relative to the Base scenario.[20] The decline in the fraction of the population for which the maternal blood MeHg levels would be

[20] Keep in mind that the Base scenario includes the SO_2 and NO_x caps of the CSA, and hence has its own reductions in Hg from FGD and SCR co-controls. The percentage change in MeHg exposure from these CSA co-controls relative to 1999 MeHg levels is estimated to rise gradually to 0.4%, or about the same magnitude as the incremental reduction shown in Table 3-6 for adding the explicit Hg provisions onto these other controls in the CSA. In other words, the Hg caps of the CSA provide only about 50% of the total estimated MeHg improvement projected for the CSA as a whole.

above the reference dose (the sum of risk reduction from both wild freshwater and marine fish) appears in the last column. Overall, the results in Table 3-6 indicate that the explicit Hg cap of the CSA would cost over $1 billion per year, and may produce an average MeHg exposure change of less than one-half of one percent. They imply expenditure of about $2.5 to 4 billion to achieve a one-percent reduction in average MeHg blood levels, implying an expenditure of about $400,000 per incrementally protected birth.

Table 3-6
National Average Cost-Effectiveness Summary for CSA Scenario
(All Changes Are Relative to the Base Scenario Outcomes)

	Incremental Cost (Millions of U.S. dollars)	Hg Emissions Change: MACT - Base (tons)		Average Hg Deposition Change		Average Change in MeHg Exposure	Change in Fraction of Population > RfD
		Elemental	Ionic	(U.S.)	(Global)		
2008	$278	-2.8	-1.2	-0.5%	-0.07%	-0.11%	-0.002%
2010	$1,129	-11.2	-4.6	-2.0%	-0.26%	-0.43%	-0.060%
2012	$1,083	-11.1	-4.2	-1.8%	-0.25%	-0.41%	-0.055%
2015	$1,450	-13.4	-4.8	-2.2%	-0.30%	-0.49%	-0.067%
2018	$2,225	-15.8	-5.4	-2.3%	-0.35%	-0.55%	-0.075%
2020	$1,425	-18.1	-3.6	-1.5%	-0.36%	-0.47%	-0.064%

3.2 Illustrative MACT

Using the same Base scenario as was used in the CSA scenario analysis, we performed a comparable set of cost-effectiveness calculations for the illustrative MACT scenario of 2.2 lb/10^{12}Btu (tBtu) implemented in 2008. Table 3-7 presents the retrofit controls selected under this scenario, and Figure 3-3 shows the emissions trends. The MACT schedule requires utilities to cut emissions over a shorter period, by 2008, pushing a majority of the total Hg-related costs to the early years of the scenario. This also eliminates much of the opportunity to benefit from the co-control of Hg that will derive from the later-imposed SO_2 and NO_x caps of the Base scenario. As a result, there are substantially more units that apply ACI than in the CSA case (173 GW compared to 122 GW, respectively). This also causes incremental costs to be much higher than for the CSA scenario, as shown in column 1 of Table 3-8, which provides the national average summary results for the illustrative MACT scenario.[21]

[21] The incremental annual costs are lower than for the CSA after 2018; this reflects differences in timing of controls. By 2020, the MACT has long been fully implemented, whereas some of the investments for the CSA Hg cap are being made in these later years. On a present-value basis, however (for 2004 through 2020), the MACT costs 3.2 times more than the CSA relative to the Base scenario (i.e., the present value of incremental costs is $19.3 billion for the MACT provisions, compared to only $6 billion for the CSA's Hg-specific provisions).

Table 3-7
Capacity (GW) Retrofitted in MACT Scenario

Year	FGD	SCR	60% Hg Control	90% Hg Control
In place as of 2000	87,612	7,891	0	0
2004	7,299	81,198		
2008	86,209	97,024	147,765	22,089
2010	1,725	2,483	2,695	
2012	110	322	223	
2015		1,410		
2018	826	1,263		
2020	3,790	7,435	21	
Total Added	99,959	191,135	150,703	22,089

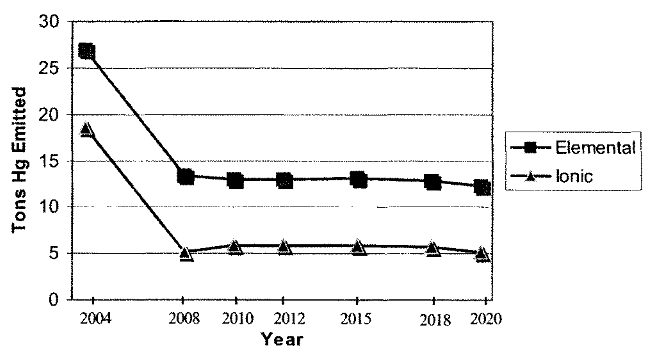


Figure 3-3
Hg^0 and Hg^{2+} Emissions Projections – MACT Case

Table 3-8
National Average Cost-Effectiveness Summary for MACT Scenario
(All Changes Are Relative to the Base Scenario Outcomes)

	Incremental Cost	Hg Emissions Change: MACT - Base (tons)		Average Hg Deposition Change		Average Change in MeHg Exposure	Change in Fraction of Population > RfD
	(Millions of U.S. dollars)	Elemental	Ionic	(U.S.)	(Global)		
2008	$4,574	-13.2	-10.0	-4.2%	-0.38%	-0.76%	-0.104%
2010	$4,016	-13.9	-8.8	-3.9%	-0.38%	-0.73%	-0.100%
2012	$3,913	-14.0	-8.5	-3.7%	-0.38%	-0.72%	-0.096%
2015	$3,275	-15.1	-7.3	-3.3%	-0.37%	-0.66%	-0.091%
2018	$2,002	-15.6	-5.8	-2.5%	-0.36%	-0.57%	-0.078%
2020	$394	-16.3	-2.8	-1.2%	-0.32%	-0.41%	-0.055%

In turn, there are earlier steps in reducing Hg deposition. There are larger reductions in Hg in the earlier years compared to the CSA, including in the Hg^{2+} species, which lowers mercury exposure from consumption of wild freshwater fish from the continental United States. Thus, Table 3-8 also shows larger average percentage changes in deposition in the earlier years than in the CSA case. Even so, MeHg exposures still fall by less than 1% on average.[22] Table 3-9 presents the estimated deposition changes within the United States by receptor site and year. Table 3-9 shows that the variation around these averages remains about the same as for the CSA scenario.

3.3 Comparing Control Policy Options

A cost-effectiveness analysis is more informative for making comparisons among alternative policy options than it is for considering any single policy scenario in isolation. We therefore now turn to a discussion of the relative cost-effectiveness of the CSA and MACT approaches to reducing U.S. electricity generation Hg emissions.

The incremental cost per reduction in births to women with MeHg blood levels above the RfD is one cost-effectiveness measure that can aid in comparisons among alternative Hg control approaches. Figure 3-4 displays this cost-effectiveness measure for the CSA and MACT for the years 2008 through 2020. It shows that the CSA approach to reducing Hg produces risk reduction at a much lower average cost – generally in the range of $400,000 per birth avoiding Hg risk. The MACT produces comparable total risk reduction earlier, but at a much greater cost

[22] Exposure changes remain less than 1% in all years even when calculated relative to 1999 levels rather than relative to the Base scenario levels, (i.e., even after adding in the effect of the co-controls in the rest of the CSA.

per unit of risk removed, starting as high as $1,000,000 per birth avoiding Hg risk in 2008. These cost-effectiveness differences are *not* a result of discounting – the annual costs are not discounted at all in this analysis. Rather, the cost-effectiveness differences occur because the MACT approach does not take advantage of any of the co-controls that can be expected from SO_2 and NO_x controls; it forces the incremental expenditure of over $4 billion per year to address Hg alone by the time of the 2008 compliance date. This expenditure is for significant amounts of ACI retrofits that do almost nothing to take advantage of future expenditures that would occur even in the Base case to reduce SO_2 and NO_x.

Table 3-9
Domestic Deposition Changes by Receptor Site (MACT compared to Base)

	2008	2010	2012	2015	2018	2020
AL00	-11.1%	-10.2%	-10.3%	-9.9%	-4.9%	0.1%
CO00	-1.2%	-2.3%	-0.8%	-1.8%	-1.6%	-0.4%
FL00	-0.9%	-0.8%	-0.8%	-0.8%	-0.5%	-0.2%
FL01	-3.7%	-3.0%	-3.1%	-1.6%	-1.0%	-0.7%
FL11	-0.3%	-0.2%	-0.2%	-0.2%	-0.1%	-0.1%
GA00	-3.9%	-3.4%	-3.4%	-3.2%	-1.9%	-0.5%
GM00	0.2%	-0.2%	-0.2%	-0.1%	-0.1%	-0.1%
LA00	-4.9%	-4.5%	-4.5%	-4.4%	-2.2%	0.0%
ME09	-0.8%	-0.8%	-0.7%	-0.7%	-0.6%	-0.5%
MI00	-3.5%	-3.5%	-3.2%	-2.7%	-2.4%	-1.3%
NH05	-5.6%	-5.6%	-4.5%	-4.3%	-3.8%	-1.6%
NJ00	-3.3%	-3.1%	-3.0%	-2.4%	-2.1%	-1.6%
NY20	-2.6%	-2.5%	-2.3%	-2.0%	-1.8%	-1.3%
PA30	-6.9%	-6.7%	-6.2%	-5.0%	-4.4%	-3.0%
TN00	-8.3%	-6.5%	-6.5%	-6.2%	-6.0%	-3.9%
TX00	-0.5%	-0.4%	-0.4%	-0.4%	-0.3%	0.0%
TX21	-12.3%	-12.0%	-12.0%	-10.6%	-8.4%	-3.3%
WI08	-2.0%	-2.0%	-1.8%	-1.6%	-1.4%	-1.0%
WI31	-7.3%	-7.3%	-6.6%	-5.5%	-4.8%	-3.1%
Average	-4.2%	-3.9%	-3.7%	-3.3%	-2.5%	-1.2%

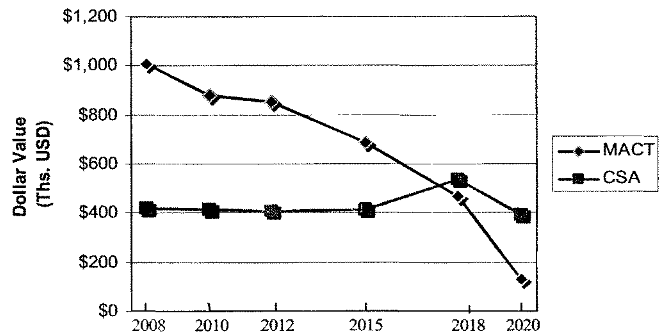


Figure 3-4
Cost per Birth Moved Below the RfD under CSA and MACT

Another comparison of alternative approaches relates to the regional differences in Hg deposition changes. By 2020, the CSA and MACT achieve similar total Hg emissions: 15 and 18 tons per year, respectively. However, this is achieved through dramatically different rules. CSA allows great flexibility regarding where the controls occur, while also allowing better leverage of SO_2 and NO_x co-controls. MACT, on the other hand, imposes rigid Hg emissions limits on each plant in the country, and is projected to be met with greater emphasis on the Hg-specific controls of ACI. Figure 3-5 contrasts the regional deposition changes of the two approaches in 2020. Each dot in Figure 3-5 represents one of the 19 receptor sites that have been modeled. The diagonal dashed line in Figure 3-5 indicates where the deposition change would be comparable in both Hg control scenarios.

This figure makes it apparent that most receptor sites experience very similar deposition reductions by 2020 under both the CSA and MACT. Two receptor sites (TX21 – Longview TX, and WI31 – Devils Lake, WI) are projected to benefit from greater deposition changes under the CSA than under the MACT in 2020. For example, the TX21 site (Longview, TX) is projected to experience over 5% reduction by 2020 in the CSA case while it would only experience about 3% reduction in the MACT case. These two areas of the country appear to benefit from the somewhat lower Hg emissions that the CSA would produce by 2020 compared to the MACT. More importantly, Figure 3-5 provides little evidence that a cap-and-trade approach will create "hot spots" compared to a MACT approach.

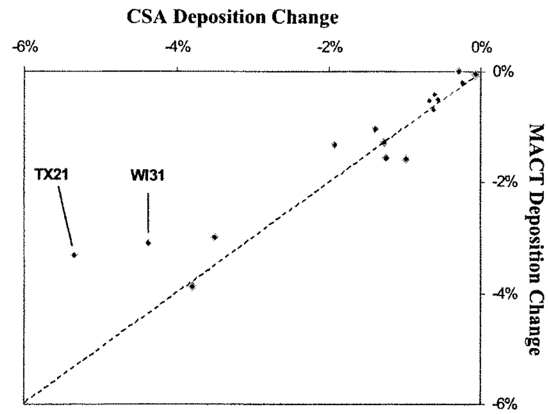


Figure 3-5
Regional Deposition Change Comparison for CSA vs. MACT, Year 2020

4
CONCLUSION

This paper has presented a framework for integrating the economics and physical sciences related to Hg emissions, fate, and deposition, and resulting exposure, to provide perspectives on the cost-effectiveness of various types of Hg policy alternatives. The key elements of the necessary calculations have been discussed, and we have presented relevant models and data for use in performing the calculations. Additionally, we have presented initial analysis results for two Hg policy proposals that are representative of the range of policy options proposed to date:

- A cap-and-trade approach integrated with a broader multi-pollutant regulatory program (based on the proposed legislation know as the Clear Skies Act), and
- A unit-specific emissions rate limitation consistent with the concept of a MACT regulation under Section 112 of the Clean Air Act.

It is important to note that the analyses presented here are preliminary and intended to illustrate the types of insights that can be derived from a comprehensive cost-effectiveness framework. More work is needed to bring this framework and its supporting data to maturity. There also remain significant gaps in the science needed to refine the framework's calculations. Any analyst considering these cost-effectiveness results will therefore need to be mindful of sensitivities and ranges of uncertainty, even as the framework is refined and its scope perhaps broadened.

Key limitations of our current framework calculations include the following:

- The estimates of reduced exposure presented here are average quantities, and do not account for geographical distributions in mercury deposition or in the consumption of wild, freshwater fish. As Tables 3-5 and 3-8 indicate, for some regions, the deposition rates can be as much as five times the average rates for the same time period. For a woman living in such a region, the effect of emissions reductions in reducing mercury exposure through the consumption of wild fresh water fish will be greater than the average.
- The distribution of MeHg exposures to U.S. women is based on the NHANES blood concentration data. However, the relative fractions of wild freshwater fish and marine fish consumed are based on average values. Women who eat more than the average fraction of wild freshwater fish will experience greater reductions in exposure than will women who eat less than the average fraction. In the limiting case, a woman who gets all her methylmercury exposure from wild fresh water fish would experience a reduced exposure proportional to the reduction in deposition in the region where she lives.
- The assumption was made that methylmercury concentrations in fish adjust instantaneously in response to changes in emissions. In actual practice, such adjustments would take time to phase in, particularly for large fish at the top of the aquatic food chain. Consequently, reductions in MeHg exposures would occur more slowly than indicated in the results presented here.

- Cost-effectiveness estimates are provided for moving a birth to a woman with blood MeHg levels above that associated with the RfD to one that is below the RfD. It is important to note that the number of women and births moving from above to below the RfD level is small because the average change in exposure is small. This measure does not recognize the full benefit of exposure reduction; the small average incremental reduction in exposure is potentially of benefit for all births to women above the RfD.
- The characterization of the benefits of MeHg exposure reductions was limited to exposure measures, e.g., the fraction of women above the RfD and the percentage change in exposure that would result from emissions reductions. No attempt was made to further characterize these exposure changes in terms of changes to the health or other characteristics of children born to women exposed above the RfD.
- EPA defines an RfD "as an estimate (with uncertainty spanning perhaps an order of magnitude) of a daily oral exposure to the human population (including sensitive subgroups) that is likely to be without an appreciable risk of deleterious effects during a lifetime." On this basis, it is not clear that any U.S. women or their children are at risk from MeHg exposure. The NHANES data indicate that about 8% of U.S. women between the ages of 16 and 49 are exposed above the RfD, but that the degree by which the RfD is exceeded is small. No woman in the NHANES data set had blood levels above the benchmark response level on which the RfD was based; only about 2.5% of U.S. women had blood levels twice those of the RfD, and the highest MeHg level measured was slightly below five times the RfD.
- The reductions in MeHg exposure under the two mercury control scenarios considered differ mainly in the speed and cost with which such reductions are achieved, with MACT reductions occurring sooner but at a higher price than CSA reductions. By 2018, the two scenarios are projected to produce similar reductions. No benefit measures were considered in this analysis that would recognize credit for earlier reductions in MeHg exposure.
- This analysis estimates the reduced exposure to MeHg that U.S. women would experience due to reduced emissions of mercury. To the extent that reductions in U.S. emissions lead to a small incremental reduction in marine fish concentrations of MeHg, exposures to MeHg also would be reduced in populations outside of the United States.

Keeping the above limitations in mind, some insights have emerged from our initial analyses that are probably fairly robust:

- The incremental reduction in the average exposure to MeHg from mercury emissions controls is small. After the reductions that are estimated to result from SO_2 and NO_x controls are considered, the additional reduction in exposures in the period out to 2020, is, on average, in the range of one half to three-quarters of a percent.
- Year 2020 MeHg exposures are lower than 1999 levels by less than 1%, under either the CSA or MACT approach.

- About half of the average U.S. exposure reduction may be associated with consumption of marine fish, rather than with consumption of freshwater fish from U.S. lakes and streams. This is due to the much greater consumption of the former, which could offset the fact that U.S. electricity generating emissions are responsible for a much smaller fraction of deposition to the world's oceans than to our inland waters.
- Cost-effectiveness can vary to a significant degree from one Hg control approach to another, even if the final emissions levels are comparable.
- The cost-effectiveness of Hg emission controls are limited in part by the significant reductions in emissions that will occur as SO_2 and NO_x controls are implemented.
- Cost-effectiveness can be greatly enhanced by coordinating Hg control requirements with reductions mandated for electricity generating emissions of SO_2 and NO_x, because these measures provide co-control of the most important part of the emissions, the ionic form of Hg.
- A cap-and-trade approach, which also can improve overall cost-effectiveness, appears not to present any significant potential to create "hot spots" compared to a MACT. In fact, it appears to offer comparable benefits deriving from utility Hg emissions reductions to all the receptor locations modeled.

Additional framework development and scenario explorations are needed as a next step. Future work that would be valuable to undertake includes:

- Developing better information on the MeHg and fish consumption patterns of the most at-risk portion of the U.S. population.
- Developing a dose-response curve across the MeHg blood level distribution.
- Incorporating better information on the geographical distribution of the wild fish that are consumed by the U.S. population.
- Developing better representation of the contribution of U.S. Hg emissions to those marine fish that are most commonly consumed by the U.S. population.
- Refining the calculations of Hg speciation by incorporating these calculations directly into the EPMM model.
- Enhancement of the atmospheric transfer coefficients to incorporate more receptor locations and greater disaggregation of source regions.

Finally, it should be noted that our use of the term "framework" rather than "model" is intentional. "Framework" represents a method of integrating information obtained from a variety of different types of models and data. We have demonstrated the use of this framework with a particular set of models, such as EPMM and TEAM. However, other models readily could be used in their place for sensitivity analyses and to build confidence that general policy insights do not derive expressly from the choice of supporting scientific model. The important element is not the particular choice of underlying model or data; it is the decision to engage in a structured, logically-consistent assessment of cost-effectiveness of alternative actions as a part of the current decision-making process.

5
REFERENCES

An Assessment of Mercury Emissions from U.S. Coal-Fired Power Plants, EPRI, Palo Alto, CA: October 2000. Technical Report #1000608.

Federal Register, 2000, "Regulatory Finding on the Emissions of Hazardous Air Pollutants from Electric Utility Steam Generating Units", *FR* 79825-79831, (December 14).

Grandjean, P., P. Weihe, R. White, et al., 1997, "Cognitive Deficit in 7-year-old Children with Prenatal Exposure to Methylmercury," *Neurotoxicol Teratol* 20:1-12.

Pai, P., et al., 1997, "Simulation of the Regional Atmospheric Transport and Fate of Mercury Using a Comprehensive Eulerian Model," *Atmospheric Environment*, 31: 17, pp. 2717-2732.

Santerre, C.R., P.B. Bush, D.H. Xu, G.W. Lewis, J.T. Davis, R.M. Grodner, R. Ingram, C.I. Wei, and J. Hinshaw, 2001, "Metal Residues in Farm-Raised Channel Catfish, Rainbow Trout and Red Swamp Crayfish from the Southern U.S.," *Food Science*, (February). See also http://news.uns.purdue.edu/UNS/html4ever/010131.Santerre.metals.html

Seigneur, C., P. Karamchandani, K. Vijayaraghavan, K. Lohman, R. Shia, 2001, "Multiscale Modeling of the Atmospheric Fate and Transport of Mercury," *Journal of Geophysical Research*, Vol. 106, No. D21, pp. 27,795 – 27,809.

Seigneur, C., K. Vijayaraghavan, K. Lohman, P. Karamchandani, C. Scott, and L. Levin, 2002, "Modeling Mercury Atmospheric Deposition in the United States," American Geophysical Union Fall Meeting, San Francisco (December).

U.S. Centers for Disease Control and Prevention, 2002, *National Health and Nutrition Examination Survey (NHANES),* National Center for Health Statistics. The data are available at http://www.cdc.gov/nchs/about/major/nhanes/NHANES99_00.htm. An early report on the first year data, "Blood and hair mercury levels in young children and women of childbearing age—United States, 1999," was published in *Morbidity and Mortality Weekly Report*, Vol. 50, No. 8, March 2, 2001.

U.S. Department of Agriculture, 2002, *Aquaculture Outlook,* Economic Research Service, LDP-AQS-15, (March 6).

U.S. EPA, 2002, *Estimated Per Capita Fish Consumption in the United States,* EPA-821-C- 02-003 (August). Available at http://www.epa.gov/waterscience/fish/consumption_report.pdf

References

U.S. EPA, 2001, *Methylmercury Assessment*, Integrated Risk Information System (CASRN 22967-92-6). Available at http://www.epa.gov/iris/subst/0073.htm.

U.S. EPA, 1997a, *Exposure Factors Handbook, Volume II, Food Ingestion Factors*, EPA/600/P-95/002Fb, (August).

U.S. EPA, 1997b, *Mercury Study Report to Congress, Volume IV: An Assessment of Exposure to Mercury in the United States,* EPA-452/R-97-006, (December).

Utility Air Regulatory Group (UARG), 2002, "UARG Comments on Process Inputs to EPA's Planned IPM Runs," submitted to EPA's Utility MACT Working Group (May 13).

A
DESCRIPTION OF THE EPMM MODEL

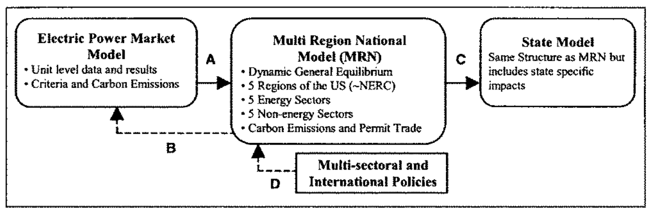
An integrated modeling framework developed by Charles River Associates, Inc. (CRA) and the Economic and Management Consulting Group (E&MC) can be used to combine a detailed simulation of electric generating unit operations and investments with simulation of how impacts of electricity sector policies would propagate to other sectors and to consumers in regional economies of the United States. This modeling framework projects emissions of four species (i.e., SO_2, NO_x, mercury from the electricity sector, and carbon emissions from all national sources). The models also explicitly simulate allowance markets subject to caps on any combination of these species. The electricity sector simulation component, known as the Electric Power Market Model (EPMM), is the focus of this section. However, we will first provide an overview of the integrated regional economic impact assessment framework of which EPMM is a core component.

A.1 Overall Analytical Approach

It is appropriate to address different policy impact questions with different types of models. Examining the impacts of sector-specific policy proposals requires a bottom-up model. Providing a full range of estimates of impacts on the economy requires a model that is more aggregated and incorporates all the important features of state, regional, or national economies. To this end, two existing models, both of which were developed and have been used successfully over a number of years, were integrated to provide economic analyses. Based on a detailed database of all power plants in the United States, E&MC Group's Electric Power Market Model (EPMM) is used to provide quantitative information about costs and changes within the electricity sector from various air emission scenarios. CRA's Multi-Region National (MRN) model is used at two stages in an analysis: (a) It provides economic and policy inputs to a detailed, plant-specific model of the electricity industry; and (b) it quantifies the economic impacts that the electricity sector has on the rest of the economy. These macro-level analyses may also be explored in more detail at the state level using the MRN State Impacts Assessment Model (SIAM).

Figure A-1 illustrates the integration of the models and the general flow of information throughout the modeling structure. For policy scenarios that are targeted on the electricity sector, we usually start the analysis using EPMM. This step produces regional information on a range of electricity sector impacts. Relevant elements of these results can then be incorporated in MRN (Linkage A in Figure A-1) to determine how fuel prices, electricity demands, and other elements of the economy (including households) would be affected. These endogenous energy responses can then be incorporated into EPMM to determine whether they would significantly change the initial results from EPMM (Linkage B). The combined results of the integrated model set provide an internally consistent picture of the economic and non-economic impacts on the economy.

If estimates of the macro-impacts are desired at the state level, the framework passes the relevant prices and constraints for each region from MRN to the state-level general equilibrium model (Linkage C in Figure A-1). Regional impacts are available from EPMM runs for 32 disaggregated regions of the United States and Canada without having to complete the latter linkage, which is only necessary in order to address impacts beyond the electricity sector.



b

Figure A-1
Integrated Framework for Multi-Pollutant Control Policy Analysis

Table A-1 lists examples of the types of impact measures that can be obtained from the integrated model set.

Table A-1
Examples of Integrated Framework Outputs

(All results available for 5-year increments from 2000-2020)
Electricity Sector Impacts **(regional disaggregation available for up to 32 NERC sub-regions)**
Retrofit decisions
Repowering decisions
Early retirements
New capacity decisions (type and amount)
Dispatch by year, season, and time-of-day (generation, capacity factors)
Wholesale electricity prices by year, season, and load time-of-day (deregulated)
Retail electricity prices by year, season, and load time-of-day (deregulated)
Capacity prices
Electricity demand changes from rest of economy (by MRN sector)
Capital expenditures
Operating expenditures
Allowance prices by pollutant
Emissions by pollutant
Fuel consumption by fuel type
Inter-regional power flows
Maintenance scheduling
Reliability (peak demand to reserve requirement ratios)
Financial value (unit, plant, company, sector)
Macro-Economic and Multi-Sectoral Impacts **(for 5 NERC-based regions in MRN/SIAM[23] and by individual state with use of SIAM; for each of the sectors in MRN[24])**
Household welfare
Gross product (national, regional, state)
Unemployment rates
Wage rates
Total energy consumption changes (natural gas, oil, coal)
Endogenous changes in energy prices (natural gas, oil, coal)
Output by sector
Labor demand by sector (% change)
Returns to shareholders
Financial value of sectors
Aggregate investment
Aggregate consumption
Consumption patterns
Vehicle miles traveled
National carbon emissions
Carbon allowance prices (for programs that extend beyond the electricity sector)
Trade flows/imports and exports (by sector, inter-regionally and internationally)

[23] The five regions currently in MRN/SIAM are: NE (NPCC & MAAC), SO (SERC & FRCC), GL (ECAR & MAIN), PL (MAPP & SPP & ERCOT), WE (WSCC) CRA has the flexibility to alter the specifications of the five regions.

[24] The sectors currently in MRN are electricity generation, energy-intensive industry, motor vehicle manufacture, other manufacturing, oil & gas extraction, coal mining, oil refining, natural gas distribution, agriculture, services. The model can also accommodate alterations in the sectoral specifications.

A.2 The Electric Power Market Model (EPMM)

EPMM was initially developed in the late 1970s to assist in assessing the implications of the proposed 1977 Amendments to the Clean Air Act. Since that time, EPMM has been expanded and updated for assessment of multi-pollutant strategies. Also, EPMM is now part of an integrated economic impact assessment framework.

EPMM employs detailed unit-level information on the more than 10,000 generating units in the United States[25] and simulates the implications of policy options on operational and generation construction decisions. The U.S. is divided into 28 regions (Figure A-2) which, for the most part, reflect North American Electric Reliability Council (NERC) regions and sub-regions. Canada is also represented with its own regions.

Environmental regulations affect decisions about:

- the mix and timing of new capacity,
- the operating life of existing fossil facilities,[26]
- the mix and timing of retrofits at existing facilities,
- fuel choice by all units,
- dispatch of all units,
- maintenance scheduling for all units, and
- the flow of power among regions.

EPMM captures all of these impacts in the process of optimizing unit responses to environmental policies.

EPMM projects these impacts through 2040.[27] EPMM determines the mix of these variables that minimizes the present value of incremental costs, which include (1) fixed and variable operating costs (including fuel costs) for all units, (2) the capital costs for investments in new plants and retrofits at existing facilities, and (3) wheeling charges.[28] This is the outcome that is expected to occur in a competitive market. In this process, EPMM also projects competitive energy prices (by season and load period) and capacity prices for each region,[29] and allowance prices for each capped emission species.

---

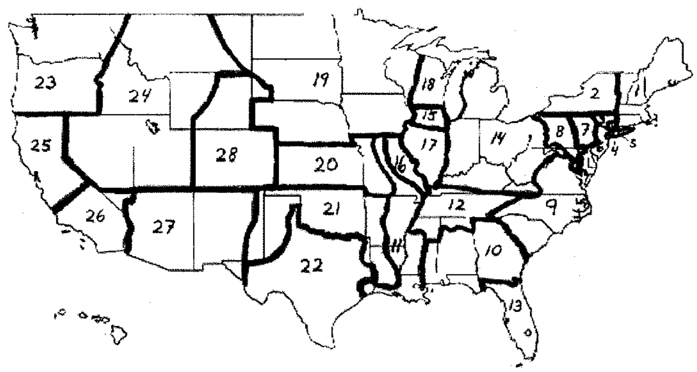
[25] The basic generating facility information is the generating unit. EPMM includes (1) units owned or recently sold by electric utilities and (2) non-utility-owned units that are connected to the grid and sell to utilities. For consistency, EPMM uses projections for most electric customers from projected demands provided by utilities. For each region, generating units are aggregated up based on prime mover, primary and secondary energy source, type of pollution control equipment, heat rate, fuel cost, and emissions rates for selected pollutants. Each unit's group is given a set of relevant retrofit control options to consider in addition to operational approaches to managing emissions control requirements.

[26] Nuclear units are treated differently from fossil units because they have licenses of fixed terms—40 years without license renewal and 60 years with license renewal.

[27] If necessary, the time horizon can be extended.

[28] These costs take into account and determine maintenance scheduling, which other models treat as fixed. The model also incorporates transmission constraints.

[29] Capacity price can vary widely by region because it depends on the mix of equipment in a region and expectations about the costs of new generating alternatives, as well as projections for future fuel prices and, particularly, natural gas prices.



1. NPCC New England	2. NPCC New York—Upstate
3. NPCC New York—Central	4. NPCC New York—Con Ed
5. NPCC New York—LIPA	6. PJM - East
7. PJM - Central	8. PJM - West
9. SERC—VACAR	10. SERC—Southern
11. SERC—Entergy	12. SERC—TVA
13. FRCC	14. ECAR
15. MAIN—Commonwealth Edison	16. MAIN—Eastern Missouri
17. MAIN—South Central Illinois	18. MAIN—Wisconsin & Upper Michigan
19. MAPP	20. SPP—North
21. SPP- South	22. ERCOT
23. WSCC—Northwest Power Pool A	24. WSCC—Northwest Power Pool B
25. WSCC—Northern California	26. WSCC—Southern California & Southern Nevada
27. WSCC—Arizona & New Mexico	28. WSCC—Rocky Mountain
29. NPCC—Maritime Provinces	30. NPCC—Hydro Québec
31. NPCC—Ontario	32. MAPP—Canada
33. WSCC—Canada	

Figure A-2
Regional Detail in EPMM

EPMM uses a linear programming algorithm to simulate a competitive market for electric power and determines, among other things, the competitive prices for energy by year, season, and time of day, and competitive capacity prices by year. A *power market* is defined as a group of regions and/or utilities interconnected by a transmission network. EPMM divides the United States and Canada into 33 interconnected power markets. The regions encompassing a market are those in which the prices are interrelated, a situation occurring when changes in circumstances in one region affect the prices for power in at least one of the other regions in the market. EPMM determines the competitive prices by balancing supplies of and demands for electricity by year, season, and by time of day for each region while taking into account transmission of electricity from one region or utility to another. The system explicitly takes into account any limits there may be on the transmission of power among regions. The demands reflect the peak demands,

energy requirements, and hourly load variations specific to each region. The supplies include all existing utility and non-utility generating units, generating facilities under construction, and potential generic new additions as well as purchase power agreements. In addition to relevant environmental regulations, for each generating facility EPMM takes into account its operating costs and characteristics, including plant type, fuel type, fuel price, heat rate, equivalent availability, equivalent forced outage rate, and annual maintenance requirements.

For each demand period within a year and over the period of analysis, EPMM orders the generating units from lowest to highest cost. It then dispatches the units to meet demands at lowest cost, taking into account the operating characteristics. In particular, EPMM meets demands in such a way as to minimize the present value of the incremental costs in the market (across all regions) for the period of analysis. The incremental costs include all fuel, other operating costs, and capital additions for nuclear plants, as well as the capital costs of both investments in new generating plants and any modifications to existing plants. EPMM does *not* explicitly take into account unamortized investments in existing generating plants because these costs cannot be affected by decisions made today or in the future. These unamortized costs are sunk, and in a competitive power market will not affect future decisions. This process provides projections of the prices that would emerge in a competitive power market because, under competition, output will be produced at lowest cost.

A.3 Projecting Competitive Energy and Capacity Prices

To understand most easily how EPMM projects energy and capacity prices, it is best to think about the process as a series of iterative steps. These steps are repeated until competitive prices are determined. *First*, the EPMM projects initial energy prices. For each of the markets, the modeling system orders the generating facilities from lowest to highest variable cost per kilowatt-hour. This variable cost has three components: (1) fuel costs; (2) emissions allowance costs where applicable (for any of SO_2, NO_x, Hg, or carbon); and (3) variable non-fuel operating and maintenance (O&M) costs where applicable. Non-dispatchable units are dispatched first to meet demands irrespective of their variable cost. Next, remaining demands are met by dispatching the lowest-cost units first and using only more expensive facilities to meet higher levels of demand. The energy price in each load period is the cost of the lowest cost unit available to meet the next increment in demand in that period. In determining when units are used to meet demands, EPMM takes into account limits on unit operations, including forced outage rates, maintenance requirements, equivalent availability, and environmental regulations.

Second, EPMM projects initial capacity prices. To provide adequate reliability within each region, systems must maintain capacity levels in excess of projected peak demands because generating units are subject to outages, and demands in peak periods can fluctuate widely. The units providing reliability are those meeting energy requirements, older facilities with relatively high variable cost per kilowatt-hour and, if a system needs new capacity, new facilities. The capacity price will be the annual incremental fixed cost per kilowatt incurred to meet the next increment in capacity requirements.

Whether a unit will (1) operate and generate electricity, (2) be used only to meet reliability requirements, or (3) be mothballed or retired depends on whether the revenues from the sale of its energy and capacity are sufficient to cover its variable and incremental fixed costs, which

include fuel, non-fuel O&M costs, and capital additions. Even if a unit is economic from the perspective of its dispatch, it will neither generate electricity nor provide reliability services *unless* the sum of its energy and capacity revenues exceeds its variable and fixed incremental costs. Otherwise, a unit will be mothballed or retired.

Nuclear plants are treated differently in making this assessment. Unlike fossil units, they cannot feasibly be removed and returned to service on a year-by-year basis. Consequently, we determine whether it is economic to continue to operate each nuclear plant over its remaining operating life (based on a current operating license of 40 years). Those plants that are found to be economic are assumed to continue to remain in service. The uneconomic nuclear plants are assumed to retire from service before 2004.

Third, EPMM looks at the opportunity to move power among markets. For each hour, power will flow from a lower-cost to a higher-cost market so long as: (1) the price differential exceeds the wheeling costs and an allowance for losses and (2) the limits on power flows between the two markets have not been reached.

Fourth, for each of the markets, EPMM determines the economically optimal mix and timing for any retrofit control options as well as new capacity additions over and above new plants already under construction. This determination will depend on two interrelated factors: (1) the need, in megawatt terms, for new capacity to meet reliability requirements; and (2) the ability of new facilities to economically displace capacity or energy from older, less efficient units. Because new capacity can be built, in part, to replace older equipment, there is an interaction in the two markets, energy and capacity. There are three cases to consider:

The first case is when existing capacity in megawatt terms is adequate to meet reliability requirements and new units cannot economically displace existing equipment. In this case, the price of capacity will be the fixed non-fuel O&M cost of the existing resource necessary to meet an increment of additional capacity requirements.

The second case is when new capacity is needed to meet reliability requirements but cannot displace energy from existing equipment. Here the price of capacity will be the annual fixed carrying charge of a new combustion turbine, the capacity type with the lowest annual fixed charge, which includes the levelized annual capital charge and the fixed non-fuel O&M costs.

The third case is when new capacity can economically displace existing equipment. Here the capacity price will be the fixed cost per kilowatt of the new capacity less the energy savings per kilowatt achieved by displacing older existing equipment. If new capacity is *not* needed to meet reliability requirements, then the difference between the fixed cost per kilowatt and the energy savings per kilowatt will be less than the fixed non-fuel O&M cost of the existing capacity that would otherwise have been used to satisfy the reliability requirements. If, on the other hand, new capacity *is* needed to meet reliability requirements, then the difference between the fixed cost per kilowatt and energy savings per kilowatt will be less than the annual levelized fixed charge for a new combustion turbine.

The process of determining (1) energy prices, (2) capacity prices, (3) power flows among markets, and (4) the timing and mix of control retrofits and/or new capacity additions is repeated until the combination is found that minimizes the present value of the *incremental* costs over the

period of the analysis for the entire market. This is the outcome that is expected in competitive power markets. The incremental costs include all of the operating costs for existing and new facilities, the capital costs of new plants and equipment, and wheeling charges. The energy and capacity prices that are consistent with the minimum present value of incremental costs are those that would emerge in competitive power markets.

A.4 Other Impact Projections

In addition to projecting competitive regional energy and capacity prices, EPMM also projects, among other things:

- The types, amounts, and timing of new capacity additions.
- Dispatch of existing and new generating plants.
- Fuel use by type for existing and new generating facilities.
- Maintenance scheduling for existing and new plants.
- Economic capacity and energy transactions among regions that are consistent with the limits imposed by transmission constraints among the regions.
- Prices of relevant emissions allowances.
- Strategies for complying with any emissions caps.

A.5 Key Assumptions

There are many assumptions underlying the projections produced by EPMM. The following includes the most important for each region:

- Peak demand and energy requirements.
- Hourly variations in electricity demand.
- Demand-side programs.
- The existing mix of utility and non-utility generating capacity.
- Current generating units under construction as well as modifications to existing facilities.
- Generic new generating options (in addition to units under construction).
- The operating costs and characteristics for existing, planned, and generic generating units:
  - Capacity
  - Plant type
  - Fuel type
  - Heat rate
  - Fuel price and characteristics
  - Non-fuel O&M costs

- Equivalent availability (for hydro and pumped storage facilities, maximum capacity factors are specified by season)
- Equivalent forced outage rate
- Maintenance requirements
- Efficiency of pumped storage facilities
- Emissions limits for SO_2, NO_x, Hg, and/or carbon
- Existing emissions control equipment for PM[30], SO_2, and NO_x.
- Percentage removal rates of emissions for existing equipment configuration
- Allocations of emissions allowances
- Plant-specific capital cost for retrofitting control equipment options for SO_2, NO_x, and Hg.
- Emissions rates for pollutants, including sulfur dioxide, nitrogen oxides, particulate matter, and carbon dioxide
- Capital and operating costs (for new generating options)

• Projections of regional fuel prices.

• Transmission-related information:
  - Limits on capacity and energy transactions among regions in a market
  - Losses for interregional transactions
  - Wheeling charges for interregional transactions

• Finance-related information:
  - Capital structure and cost of money
  - Income tax rates
  - Property tax and insurance rates
  - Book life for new generating options
  - Tax life for new generating options
  - Treatment of deferred taxes
  - Construction period for new generating options

[30] The PM control equipment is relevant information for estimating initial Hg control levels, with differing rates of Hg control depending on the type of PM equipment.

A.6 Data Sources

EPMM uses data from the credible sources, and its data have been reviewed by utilities over years of model development, but also by many individual utilities as part of a major, on-going utility industry analysis project.

Electricity Market-Related Data. EPMM uses Energy Information Administration (EIA) and Federal Energy Regulatory Commission (FERC) data on unit location, age, size, heat rate, and current fuels to assess the marketability of each unit's power at different points in the load curve. Delivered coal prices are based on FERC-reported data. EIA's *Annual Energy Outlook 2003* data are used for other fuel prices and for fuel price trends. Electricity demand forecasts come from NERC. EPMM endogenously generates a forecast of new capacity that minimizes the costs of meeting these other forecasts while accounting for the economics of the existing units in each model region.

Emissions-Related Data. For existing unit emissions of SO_2 and NO_x, EPMM contains EPA's Continuous Emission Monitoring Systems (CEMS) data. Carbon emissions are calculated using the same carbon contents for each fossil fuel used for the U.S. Greenhouse Gas Emissions Inventory. Hg emissions require more complex sets of inputs. The Hg contents of various coals are from EPRI's estimates of unit Hg emissions, based on the 1999-2000 ICR data. However, unit Hg emissions are equal to the Hg in the coal times an emission reduction factor that was based on a 2001 utility industry expert consensus. These factors depend on the existing PM, SO_2 and NO_x control equipment, and on the rank of the coal. They were founded on the EPRI analysis of EPA's ICR data, but have been adjusted by the group of experts based on best judgment to reflect new information not represented in the ICR data. The majority of these adjustments relate to the co-control that an SCR adds. At present, SCRs are assumed to provide co-control only if the coal being burned is a non-western bituminous coal. SNCRs are not assumed to offer any co-control of Hg.

EPMM contains options for retrofits of control technologies that reduce emissions. Assumptions for the cost and percent emissions reduction of control retrofits for SO_2 and NO_x controls are taken from assumptions used by EPA in their own emissions model exercises for multi-pollutant bills. The model contains options to add a wet FGD, SCR, Low-NO_x Burner (LNB), and sometimes Selective Non-Catalytic Reduction (SNCR). These technologies may also reduce Hg, following the same co-control assumptions used to estimate current Hg emissions.

Hg control technology data are far more subject to judgment, and the EPMM assumptions are based on a view of the costs and potential effectiveness developed by EPRI's Hg control technology research staff. These assumptions are expected to be updated relatively frequently as Hg control technology evolves. The table of co-control assumptions is also likely to evolve more rapidly than other control technology assumptions in EPMM. Currently EPMM contains a set of Hg control options that are all based on activated carbon injection (ACI). If the unit considering the retrofit has a fabric filter in place, then it can adopt a simple ACI. All other units must also apply a polishing fabric filter in conjunction with the ACI. All units choosing to add a Hg control retrofit can also choose to operate the equipment to achieve a 60% incremental removal of Hg or a 90% total removal of Hg. The lower removal efficiency is achieved with one-fourth the carbon injection rate as the higher removal efficiency, and hence at a lower operating cost. Otherwise the equipment (i.e., its capital cost) is the same.

B

DETAILS OF CALCULATIONS TO SPECIATE EPMM'S TOTAL HG EMISSIONS PROJECTIONS

The EPRI report "An Assessment of Mercury Emissions from U.S. Coal-Fired Power Plants" (EPRI, 2000) provides in its Table 3-2 to 3-7 a set of equations that predict the percent removal of Hg and the percent of Hg emitted as Hg^0 for different configurations of PM and SO_2 control equipment. The relationships for overall percent removal were used as a foundation for the co-control assumptions of EPMM. However, because EPMM must address co-controls from future NO_x controls as well as from PM and SO_2 equipment, the percent removal assumptions also required expert judgment. Table B-1 shows the overall set of total percent removal assumptions that were derived by a group of industry experts relying on information such as that of EPRI's Table 3-2 to 3-7, but also relying on more recent experiences in field studies (UARG, 2000). These are the percent removal assumptions used in EPMM, often referred to as the "industry consensus co-control assumptions."[31]

The extra step that is needed for the cost-effectiveness framework is to formulate a comparable set of assumptions for the percent of the remaining emitted Hg that is emitted as Hg^0, Hg^{2+}, and Hg_p. For this, we returned to the original foundation for EPMM's percent removal assumptions – the EPRI Table 3-2 to 3-7 regression results. These relationships are a function of the chlorine contents of the fuels being burned. EPMM does not have information on the chlorine content of the coals being burned in the model, but it does provide information on the rank of coals being burned. Although the chlorine content varies from coal to coal, there is a pronounced difference in the average levels of different ranks of coals. To obtain general estimates for use in the current framework speciation calculations, we assigned a chlorine content of 1000 ppm to the average eastern bituminous coal, and a chlorine content of 100 ppm to the subbituminous, lignite and western bituminous coals.[32] When these values are inserted into the relationships derived by EPRI, one derives the estimates of the percent of Hg emitted as Hg^0 for all of the "no SCR" cases.

[31] The group of experts felt that SNCR would not reliably produce incremental Hg reductions, so the only NO_x control shown in Table B-1 is SCR.

[32] These values were derived by a review of the distribution of chlorine contents in coals of each rank, using EPRI's data on coals burned by each unit. Actual levels vary widely within each rank.

Table B-1
"Industry Consensus" on Hg Co-Control Assumptions

PM Equipment	SO_2 Equipment	NO_x Equipment	% Removal of Hg		
			Bituminous	Subbituminous	Lignite
FF	Dry FGD	No SCR	83	15	8
		SCR	90	15	8
	Wet FGD	No SCR	85	76	12
		SCR	90	76	12
	None	No SCR	74	65	4
		SCR	74	65	4
CSESP	Dry FGD	No SCR	50	25	20
		SCR	75	25	20
	Wet FGD	No SCR	59	33	42
		SCR	75	33	42
	None	No SCR	36	20	7
		SCR	36	20	7
HSESP	Dry FGD	No SCR	n/a	n/a	n/a
		SCR	n/a	n/a	n/a
	Wet FGD	No SCR	55	13	10
		SCR	75	13	10
	None	No SCR	18	2	0
		SCR	18	2	0
Venturi	Dry FGD	No SCR	n/a	n/a	n/a
		SCR	n/a	n/a	n/a
	Wet FGD	No SCR	24	4	12
		SCR	40	4	12
	None	No SCR	20	2	5
		SCR	20	2	5

The remaining step needed was to assess the impact on speciation of adding an SCR to each of the "no SCR" configurations. To do this, we built a structural representation of the Hg speciation as it flows from the boiler exit to the stack, accounting for the fact that each type of PM control equipment would remove effectively all of the Hg_p that has formed by the time it enters the collector device, and that a wet FGD would remove effectively all of the Hg^{2+} that has formed.

The parameters for this structural model were calibrated to the data in EPRI's Table 3-2 to 3-7 for each rank of coal. Once the structural model was calibrated to all of the no-SCR configurations, we then introduced an SCR into the model. The SCR is assumed to cause an additional fraction, F, of the Hg^0 to be converted to Hg^{2+} as it passes through the device. This shift causes a greater share of gases to be captured when they subsequently pass through other pollution control devices on the unit. The additional Hg_p that is formed from the additional gaseous Hg^{2+} would be removed from the PM collector. All remaining additional gaseous Hg^{2+} would be captured by a wet FGD. We iterated to find a value for F that would produce SCR Hg removal co-benefits with the same pattern and magnitude that is observed in Table B-1. Once that value of F was fixed, the model allows us to derive a set of percent Hg^0 values for the "with SCR" configurations that is internally consistent with to set of Hg removal co-benefits to which it was calibrated.

Overall, this structural model did a very good job at producing the SCR co-benefits pattern in Table B-1 when we used F=50% for non-western bituminous coals. This suggests that the industry consensus SCR co-benefits assumptions imply that an SCR causes about 50% of the erstwhile Hg^0 in the flue gases to be converted to Hg^{2+} in high-chlorine coals. Because the industry consensus co-benefits assumptions currently assume zero co-benefits for lower rank and western bituminous coals, this implies F=0% for these other coals.[33]

The results are summarized in Table B-2, showing the percent of emissions that are in the form of Hg^0 from any unit as a function rank of the coal it is burning, and the emissions control equipment in place on that unit. The framework thus uses those additional EPMM outputs to estimate the Hg^0 emissions in each of the 28 EPMM regions, following the assumptions shown in Table B-2. Hg^{2+} emissions are considered to make up the remainder of the total Hg emitted, and we assume that Hg_p remains a very minor, constant portion of the emissions. This is done at the regional level.

EPMM incorporates two types of retrofits that are designed specifically to reduce Hg emissions: Activated Carbon Injection (ACI) with and without a polishing fabric filter. ACI alone is used if a plant has a fabric filter, and the polishing fabric filter is applied on all other types of units. We assume that both of these retrofits work in conjunction with the other emissions control equipment to eliminate a maximum of 90% of elemental and ionic Hg emissions. Thus, they affect the percent of Hg removed, but preserve the speciation that existed prior to their installation.

Due to limited data, there is much uncertainty about the effects of different emissions control equipment combinations on Hg emissions. Several characteristics of electricity generating units and their fuel, ranging from actual coal chlorine content to heat rate to fly ash concentration, will cause variability in their actual outcomes. However, at present these relationships are the best available, given data limitations, and the limitations of the science as well.

---

[33] However, we noted that using F=50% for these coals too implies some but much smaller Hg removal co-benefits for the lower chlorine coals. Although the resulting SCR co-benefits are non-zero, they are so close to the Hg removals of the no-SCR case that they might not have been observed in the field.

Table B-2
Hg Emissions Speciation Assumptions

PM Equipment	SO_2 Equipment	NO_x Equipment	Percent Hg Emitted As Hg^0		
			Bituminous	Subbituminous	Lignite
FF	Dry FGD	No SCR	70	96	96
		SCR	63	96	96
	Wet FGD	No SCR	78	94	94
		SCR	63	94	94
	None	No SCR	6	33	33
		SCR	4	33	33
CSESP	Dry FGD	No SCR	91	96	96
		SCR	86	96	96
	Wet FGD	No SCR	94	98	98
		SCR	86	98	98
	None	No SCR	34	63	63
		SCR	18	63	63
HSESP	Dry FGD	No SCR	n/a	n/a	n/a
		SCR	n/a	n/a	n/a
	Wet FGD	No SCR	95	99	99
		SCR	89	99	99
	None	No SCR	42	79	79
		SCR	22	79	79
Venturi	Dry FGD	No SCR	n/a	n/a	n/a
		SCR	n/a	n/a	n/a
	Wet FGD	No SCR	30	70	70
		SCR	30	70	70
	None	No SCR	30	70	70
		SCR	30	70	70

C

DETAILS OF RECEPTOR LOCATIONS FOR TRANSFER COEFFICIENTS BASED ON AER'S TEAM MODEL

Receptor Label	Receptor	MDN site or location	Latitude	Longitude	TEAM grid cell
WI08	Brule River, WI	WI08	46° 45' 00" N	91° 30' 00" W	34, 29
WI31	Devil's Lake, WI	WI31	43° 25' 05" N	89° 43' 55" W	35, 25
MI00	Lake Michigan eastern shore, MI	Sleeping Bear Dunes National Lakeshore			38, 27
PA30	Lake Erie, PA	PA30	42° 09' 21" N	80° 06' 48" W	44, 25
NY20	Huntington Wildlife Refuge, NY	NY20	43° 58' 21" N	74° 13' 15" W	49, 29
ME09	Greenville Station, ME	ME09	43° 29' 23" N	69° 39' 52" W	51, 32
NH05	New Castle, NH	NH05	43° 10' 00" N	70° 52' 00" W	52, 29
NJ00	Pines Lake, NJ	Passaic County	40° 59' 32" N	74° 16' 06" W	50, 25
CO00	McPhee/Narraguinnep Reservoirs, CO	Southwestern Colorado	37° 34' 37" N	94° 42' 42" W	31, 11
TN00	Great Smoky Mountains National Park, TN	Eastern Tennessee			43, 16
TX21	Longview, TX	TX21	32° 22' 43" N	94° 42' 42" N	31,11
TX00	Upper Lavaca Bay, TX	Near Point Comfort, Calhoun County	28° 39' 35" N	96° 35' 40" W	29, 5
LA00	Louisiana/Mississippi southern border				38, 8
AL00	Mobile Bay, AL	South coast Alabama			39, 9
GA00	Ichawahoa-chaway Lake, GA	Southwest GA: 30 miles west of Albany	31.6° N	84.6° W	43, 11
FL00	Apalachicola Bay, FL	Florida panhandle			42, 8
FL01	Lake Barco, FL		29° 40' 32" N	82° 00' 31" W	46, 9
FL11	Everglades National Park, FL	FL11	25° 23' 24" N	80° 40' 48" W	49, 3
GM00	Gulf of Mexico				37, 2

D

BLOOD LEVELS OF METHYLMERCURY AND FISH CONSUMPTION: INFORMATION FROM THE NHANES 2002 DATA RELEASE

D.1 Introduction

Data from the federal National Health and Nutrition Examination Survey (NHANES), released in June 2002, provided the first opportunity to compare blood levels of total mercury with the amount of fish consumed for a large number of U.S. citizens. This comparison is important because people are exposed to mercury mainly by eating fish that have accumulated MeHg in their flesh. The NHANES data make this comparison possible for a large number of U.S. women and children who have participated in the survey.

Designed to collect information about the health and diet of people in the United States, NHANES has been conducted for more than 40 years by the National Center for Health Statistics, part of the Centers for Disease Control and Prevention (CDC). The most recent survey, begun as an ongoing study in mid-1998, examines about 5,000 people each year. First-year results (1999 data files) from this survey reported measurements of mercury in the blood and hair of women and children.[34] In June 2002, the results of many of the 1999-2000 data files were released, including measurements of mercury in blood and the responses to dietary questionnaires that asked people about their fish consumption. In March 2003, the full set of data on hair mercury concentrations was released.

The data made available in June 2002 include the results of measurements of concentrations of mercury in blood for 1,709 women between the ages of 16 and 49, and for 705 children between the ages of 1 and 5. Weighting factors are provided so that results from the sample population can be adjusted to reflect the age, geographic location, and ethnicity of the U.S. population as a whole.

D.2 Blood Concentrations

Of the 1,709 women whose blood levels of mercury were measured, 286 were pregnant. Seven more pregnant women participated in the study, but their blood levels of mercury were not measured and their ages fell outside the 16-49 range. Thus, using the 16-49 age group did not eliminate any pregnant women for whom blood mercury data were available.

[34] Centers for Disease Control and Prevention, "Blood and hair mercury levels in young children and women of childbearing age—United States, 1999," *Morbidity and Mortality Weekly Report*, Vol. 50, No. 8, March 2, 2001.

The most recent results report both total and speciated mercury in blood. According to documentation accompanying release of the data files, "speciated" mercury refers to inorganic (elemental and ionic) mercury. Thus, total mercury includes these inorganic forms plus organic methylmercury. The NHANES data on blood levels of speciated mercury appear to be unusable, except for the upper tail of the distribution. For 97% of all samples the reported blood levels of speciated (i.e., inorganic) mercury were either 0.3 or 0.4 µg/L, with the lowest reported value (or detection limit divided by the square root of two) being 0.3 µg/L. For about 25% of these samples, reported blood levels of speciated mercury exceeded blood levels of total mercury, which could be as low as 0.1 µg/L. Clearly, the high detection limits for speciated mercury make it impossible to determine what fraction of total mercury is due to speciated (inorganic) forms. At the upper end of the total mercury distribution, blood levels were sufficiently high that a meaningful estimate could be made of the distribution of total mercury minus inorganic mercury. However, the problem of high detection limits for inorganic mercury does not permit the calculation of the mean blood concentration of methylmercury.

The data released in 2002 show lower blood levels of total mercury, and suggest lower mercury exposures, than the first-year data reported in 2001. Table D-1 and Figure D-1 describe the distribution of the June 2002 results. When the data released in 2002 are weighted to account for U.S. demographic patterns, about 9% of women in the study have blood levels of total mercury above 5.37 ppb. Since the reference dose (RfD) is for methylmercury, not for total mercury, the number of women exposed above the RfD for methylmercury is lower.

Table D-1
Blood Mercury Levels of Women Ages 16–49

	All Women in Sample	Pregnant Women
Number of women in sample	1709	286
Geo. mean, total blood Hg	1.02*, 0.92† ppb	0.97*, 0.78† ppb
Arithmetic mean, total blood Hg	2.00*, 1.72† ppb	1.86*, 1.45† ppb
90th percentile	4.9*, 4.0† ppb	5.0*, 3.1† ppb
75th percentile	2.1*, 1.9† ppb	2.0*, 1.6† ppb
50th percentile	1.0*, 0.9† ppb	1.0*, 0.8† ppb
25th percentile	0.5*, 0.4† ppb	0.4*† ppb
10th percentile	0.2*† ppb	0.2*† ppb
Exceed RfD blood level of 5.37 ppb§	8.7%*, 6.0%†	9%*, 4.9%†

* Weighted for demographic factors
† Unweighted for demographic factors
§ This compares total mercury blood concentrations with the reference dose (RfD) for methylmercury.

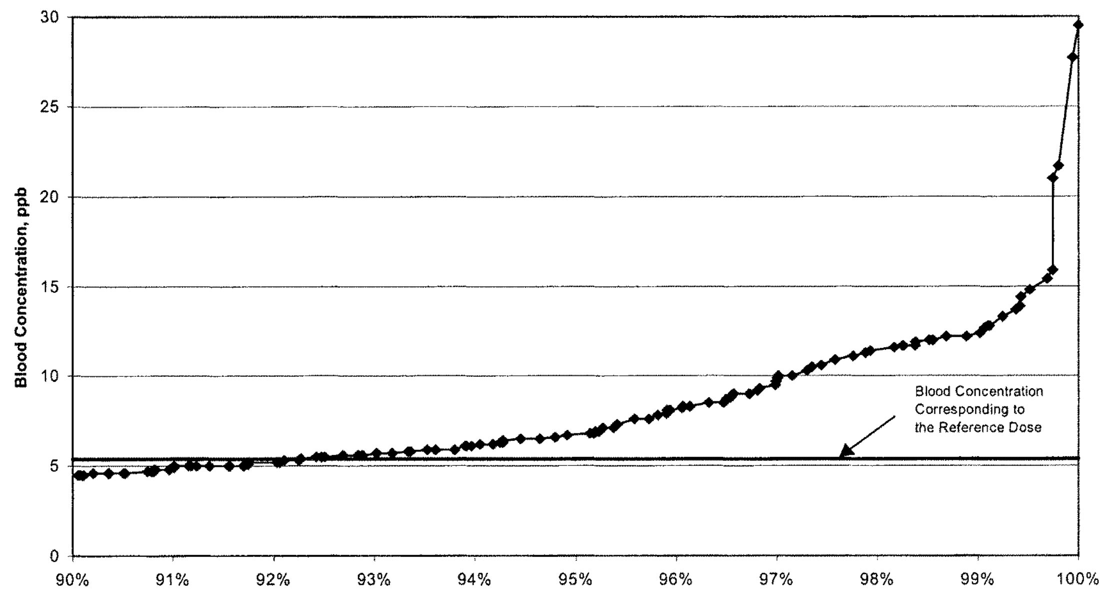


Figure D-1
Distribution of Top 10% of MeHg Blood Levels, U.S. Women, Ages 16-49, with Demographic Weights Applied

As noted above, the upper tail of the distribution for methylmercury can be calculated, because for the women with the highest blood mercury concentrations, the imprecision in the inorganic mercury measurements do not greatly affect the difference between total mercury and inorganic mercury. When this methylmercury distribution is examined, the data indicated that slightly over 5% of the women in the sample, before demographic weights are applied, have blood methylmercury concentrations above the RfD. After applying the demographic weights, it appears that about 7.7% of the women have blood concentrations above the RfD.

D.3 Fish Consumption

Data on fish consumption are available for 1,646 of the 1,709 women whose blood levels of mercury were measured, for 277 of the 286 pregnant women in that group, and for 667 of the 705 children. The NHANES data include detailed information on food consumed within the previous 24 hours, and a count of how many times various types of fish were consumed over the previous 30 days. Types of fish in the NHANES dataset include: breaded fish products, bass, catfish, cod, flatfish, haddock, mackerel, perch, pike, pollock, porgy, salmon, sardines, sea bass, shark, swordfish, trout, tuna, walleye, "other" fish, and "unknown" fish. Information about types of fish consumed is potentially useful for determining patterns of fish consumption among women with the highest measured blood levels of mercury. Conversely, this information can also be used to compare the blood levels of mercury in women who consume fish far more frequently than average.

D.4 Fish Consumption versus Mercury Concentration

Table D-2 shows the number of fish meals women consumed in the previous 30 days when the women are grouped according to their measured blood levels of total mercury (expressed as blood mercury percentiles). It includes only those women for whom both fish consumption and blood level data are available. From the table, it is clear that blood levels of mercury rise as fish consumption rises. This relationship holds for women in the 16–49 age group, and for pregnant women within that group. It is noteworthy that the woman with the highest blood level of mercury (measured at 38.9 ppb, see Table 3-5) reported eating no fish during the previous 30 days.

Figure D-2 also illustrates the positive trend observed when the number of fish meals women consumed is compared with their measured blood levels of total mercury. The line on the figure represents the best linear fit to the data. However, it is clear that the data points are widely scattered and that accurate prediction of an individual's blood level of mercury based on her fish consumption is impossible.

Tables D-3 and D-4 show the number of fish meals reported over the previous 30 days for each type of fish. It is important to know how many freshwater fish women in the survey consumed because mercury emissions from U.S. utilities are likely to have a much greater impact on the mercury content within freshwater fish than marine fish. However, it is difficult to get a clear picture of freshwater fish consumption by women in the survey for two reasons. First, the total number of meals reported for "other" or "unknown" fish exceeds the number of meals reported for freshwater fish. Until it is known how many of these other and unknown fish are freshwater species, the actual number of freshwater fish meals consumed by women in the survey remains uncertain.

Second, catfish and trout make up almost 80% of the freshwater fish meals reported for all women, and more than 90% of such meals for pregnant women and children. Unfortunately, it is impossible to tell from the data whether these catfish and trout meals involve wild or farm-raised fish. The distinction is important because farm-raised catfish and trout have lower methylmercury levels than their wild cousins (Santerre, 2001). This is so because farm-raised fish are housed and cultivated using specialized feed in a way that prevents methylmercury from accumulating in their bodies, as it does in wild fish. For this reason, mercury exposures from consumption of farm-raised catfish and trout are already low, and are unlikely to be affected by changes in U.S. utility mercury emissions.

As Table D-4 indicates, the fraction of fish consumed that are wild, freshwater fish is not clear from the data. The fraction depends on the portion of catfish, trout, other, and unknown fish that are wild freshwater fish. If we assume that most of the trout and catfish are farm-raised, that some of the breaded fish is catfish, and that about half of the salmon is farm-raised, then the total number of fish meals consumed by women ages 16-49 in this sample would be of the order of 3,000 fish meals. Of this number, 84 fish meals are of species that are almost certain to be wild, freshwater fish, and another 500 or so meals are of "other" and "unknown" fish. On this basis, we estimate that the fraction of fish meals eaten that are of wild, freshwater fish, is not likely to exceed 10% of the total when meals of farm-raised fish are excluded. This estimate is probably high, but a high estimate was made to avoid underestimating the fraction of wild freshwater fish meals because the effect of mercury emission reductions on the methylmercury content of fish is greatest for such fish. Having selected an estimate of 10% as the fraction of fish meals, excluding farm-raised fish, that are wild freshwater fish, then we will assume that the remaining 90% of fish meals are of marine fish.

Table D-2
Fish Consumption (Meals in the Previous 30 Days) Versus Blood Mercury Levels (Unweighted for Demographic Factors)

Blood Hg Percentile	Women 16–49	Pregnant Women, 16–49
All	2.15	2.16
Top 90th percentile	5.41	7.46
Top 75th percentile	4.08	4.57
Top half of group	3.25	3.27
Lowest 25th percentile	0.73	0.72
Lowest 10th percentile	0.54	0.51
Those with blood Hg > RfD level*	6.22	8.92

*Refers to women with blood levels of total mercury exceeding the reference dose (RfD) for methylmercury.

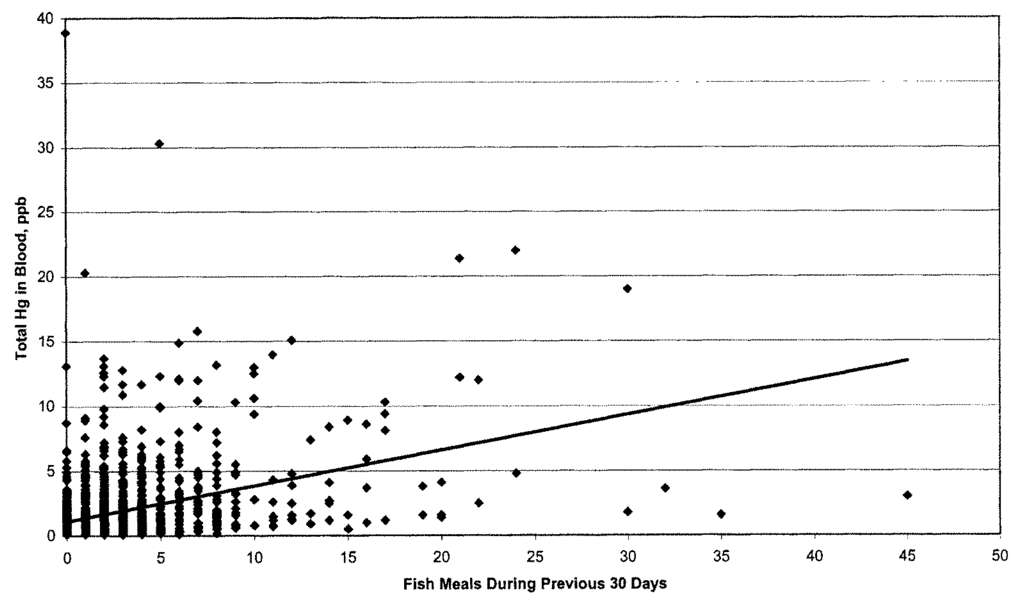


Figure D-2
Blood Mercury Versus Fish Meals Women Ages 16-49

Table D-3
Fish Meals Reported for the Previous 30 Days by Women Ages 16–49 and Children Ages 1-5

Fish Type	Meals Reported by all Women (1646 in sample)	Meals Reported by Pregnant Women (277 in sample)	Meals Reported by Children (667 in sample)
Breaded fish	403	48	274
Bass	24	1	5
Catfish	253	44	81
Cod	105	13	17
Flatfish	92	11	51
Haddock	38	6	13
Mackerel	18	1	7
Perch	46	4	0
Pike	1	0	2
Pollock	67	8	9
Porgy	9	0	1
Salmon	341	44	72
Sardines	56	8	24
Sea bass	15	4	2
Shark	4	2	1
Swordfish	28	12	0
Trout	49	8	13
Tuna	1,470	289	460
Walleye	11	0	3
Other fish	385	58	154
Unknown fish	126	37	42
Total fish meals	3,541	598	1,231

Table D-4
Summary of NHANES Data on Fish Meals

Fish Type	Meals Reported by all Women (1646 in sample)	Meals Reported by Pregnant Women (277 in sample)	Meals Reported by Children (667 in sample)
Total fish meals	3,541	598	1,231
Total freshwater fish*	384	57	104
Total freshwater fish other than trout and catfish*	82	5	10
Catfish and trout meals	302	52	94
Total "other" and "unknown" fish meals	511	95	196

* If "breaded," "other," and "unknown" fish are not included as freshwater fish.

D.5 Summary

In summary, the NHANES data released in 2002 provide the first opportunity to compare a biomarker for mercury exposure—blood level of total mercury—with information about fish consumption for a large number of individuals. Although the data show that blood levels of mercury rise as fish consumption rises for the sample population as a whole, data variability precludes accurate prediction of an individual's blood level of mercury based on her fish consumption.

When the data are weighted to account for U.S. demographic patterns, about 7.7% of women in the study have blood levels of methylmercury above 5.37 ppb, blood concentration associated with the reference dose (RfD).

As noted, U.S. utility emissions of mercury would contribute to these exposures primarily by adding mercury to the loading in freshwater fish that are then consumed. About 2.3% of fish meals eaten by women in the survey were of wild freshwater fish, and 10–25% were of freshwater fish that may be either wild or farm-raised, depending on assumptions made about the undefined species in the survey. Based on the data and explanation provided above, it appears that for fish types that would be affected by reductions in mercury emissions (that is, all fish except farm-raised fish), 10% is a conservatively high estimate of the fraction that are wild, freshwater fish.

Program:
Air Toxics Health and Risk Assessment

About EPRI

EPRI creates science and technology solutions for the global energy and energy services industry. U.S. electric utilities established the Electric Power Research Institute in 1973 as a nonprofit research consortium for the benefit of utility members, their customers, and society. Now known simply as EPRI, the company provides a wide range of innovative products and services to more than 1000 energy-related organizations in 40 countries. EPRI's multidisciplinary team of scientists and engineers draws on a worldwide network of technical and business expertise to help solve today's toughest energy and environmental problems.

EPRI. Electrify the World

STATEMENT OF DEBORAH C. RICE, PH.D., MAINE DEPARTMENT OF ENVIRONMENTAL PROTECTION, AUGUSTA, MAINE

I would like to thank the Committee for this opportunity to present information on the adverse health consequences of exposure to methylmercury in the United States. Until 3 months ago, I was a senior toxicologist in the National Center for Environmental Assessment in the Office of Research and Development at the Environmental Protection Agency. I am a co-author of the document that reviewed the scientific evidence on the health effects of methylmercury for EPA, and which included the derivation of the acceptable daily intake level for methylmercury.

I would like to focus my presentation on four points that are key to understanding the health-related consequences of environmental mercury exposure. One: there is unequivocal evidence that methylmercury harms the developing human brain. Two: the Environmental Protection Agency used analyses of three large studies in its derivation of an acceptable daily intake level, including the study in the Seychelles Islands which found no adverse effects. Three: 8 percent of women of child-bearing age in the United States have levels of methylmercury in their bodies above this acceptable level. And four: cardiovascular disease in men related to low levels of methylmercury has been documented, suggesting that a potentially large segment of the population may be at risk for adverse health effects.

The adverse health consequences to the nervous system of methylmercury exposure in humans were recognized in the 1950's with the tragic episode of poisoning in Minamata Bay in Japan, in which it also became clear that the fetus was more sensitive to the neurotoxic effects of methylmercury than was the adult. A similar pattern of damage was apparent in subsequent episodes of poisoning in Japan and Iraq. These observations focused the research community on the question of whether exposure to concentrations of methylmercury present in the environment might be producing neurotoxic effects that were not clinically apparent.

As a result, over half a dozen studies have been performed around the world to explore the effects of environmental methylmercury intake on the development of the child. Studies in the Philippines (Ramirez *et al.*, 2003), the Canadian Arctic (McKeown-Eyssen *et al.*, 1983), Ecuador (Counter *et al.*, 1998), Brazil (Grandjean *et al.*, 1999), French Guiana (Cordier *et al.*, 1999) and Madeira (Murata *et al.*, 1999) all found adverse effects related to the methylmercury levels in the children's bodies. These included auditory and visual effects, memory deficits, deficits in visuospatial ability, and changes in motor function.

In addition to the above studies, there have been three major longitudinal studies on the effects of exposure to the mother on the neuropsychological function of the child: in the Faroe Islands in the North Atlantic (Grandjean *et al.*, 1997), in the Seychelles Islands in the Indian Ocean (Myers *et al.*, 1995), and in New Zealand (Kjellstrom *et al.*, 1989). Two of these studies identified adverse effects associated with methylmercury exposure, whereas the Seychelles Islands study did not. Impairment included decreased IQ and deficits in memory, language processing, attention, and fine motor coordination. A National Research Council (NRC) National Academy of Sciences panel evaluated all three studies in their expert review, concluding that all three studies were well designed and executed (NRC, 2000). They modeled the relationship between the amount of methylmercury in the mother's body and the performance of the child on a number of neuropsychological tests. From this analysis, they calculated a defined adverse effect level from several types of behavior in each of the three studies. These adverse effect levels represent a doubling of the number of children that would perform in the abnormally low range of function. The National Academy of Sciences panel also calculated an overall adverse effect level of methylmercury in the mother's body for all three of the studies combined, including the negative Seychelles study. Thus the results of all three studies were included in a quantitative manner by the NRC.

The Environmental Protection Agency (EPA) used the analyses of the NRC in the derivation of the reference dose, or RfD, for methylmercury. The RfD is a daily intake level designed to be without deleterious effects over a lifetime. The EPA divided the defined deleterious effect levels calculated by the NRC by a factor of 10 in its RfD derivation. There are two points that need to be made in this regard. First, the factor of 10 does not represent a safety factor of 10, since the starting point was a level that doubled the number of low-performing children. Second, the EPA performed the relevant calculations for a number of measurements for each of the two studies that found deleterious effects as well as the integrative analysis that included all three studies modeled by the NRC, including the negative Seychelles study. The RfD is 0.10 μg/kg/day based on the Faroe Islands study alone or the integrative analysis of all three studies. The RfD would be lower than 0.10 μg/kg/day if only the New Zealand study had been considered. Only if the negative Seychelles

Islands study were used exclusively for the derivation of the RfD, while ignoring the values calculated for the Faroe Islands and New Zealand studies, would the RfD be higher than the current value of 0.10 µg/kg/day. EPA believes that to do so would be scientifically unsound, and would provide insufficient protection to the U.S. population.

A substantial portion of U.S. women of reproductive age have methylmercury in their bodies that is above the level that corresponds to the EPA's RfD. Data collected over the last 2 years as part of the National Health and Nutritional Examination Survey (NHANES 99+) designed to represent the U.S. population (CDC, Web) revealed that about 8 percent of women of child-bearing age had blood levels of methylmercury above the level that the U.S. EPA believes is "safe" (Schober *et al.*, 2003). This translates into over 300,000 newborns per year potentially at risk for adverse effects on intelligence and memory, ability to pay attention, ability to use language, and other skills that are important for success in our highly technological society.

I would like to further comment here on the use of a factor of 10 by EPA to derive the allowable daily intake level (RfD) for methylmercury from the defined effect levels calculated by the National Research Council. The RfD corresponds to roughly 1 part per million (ppm) of methylmercury in maternal hair, from the defined effect level of about 11 ppm calculated by the NRC. But we know that there is no evidence of a threshold below which there are no adverse effects down to about 2–3 ppm in hair, the lowest levels in the Faroe Islands study. In fact, there is evidence from both the Faroe Islands (Budtz-Jorgensen *et al.*, 2000) and New Zealand (Louise Ryan, Harvard University, personal communication) studies that the change in adverse effect in the child as a function of maternal methylmercury level may be greater at lower maternal methylmercury levels than at higher ones. Therefore, the so-called safety factor almost certainly is less than 10, and may be closer to nonexistent. Babies born to women above the RfD may be at actual risk, and not exposed to a level 10 times below a risk level.

There is an additional concern regarding the potential for adverse health consequences as a result of environmental exposure to methylmercury. Several years ago, a study in Finnish men who ate fish found an association between increased methylmercury levels in hair and atherosclerosis, heart attacks, and death (Salonen *et al.*, 1995, 2000). Two new studies in the U.S. and Europe found similar associations between increased methylmercury levels in the bodies of men and cardiovascular disease (Guallar *et al.*, 2002; Yoshizawa *et al.*, 2002). Effects have been identified at hair mercury levels below 3 ppm. It is not known whether there is a level of methylmercury exposure that will not cause adverse effects. It is important to understand that the cardiovascular effects associated with methylmercury may put an additional, very large proportion of the population at risk for adverse health consequences as a result of exposure to methylmercury from environmental sources.

In summary, there are four points that I would like the Committee to keep in mind. First, at least eight studies have found an association between methylmercury levels and impaired neuropsychological performance in the child. The Seychelles Islands study is anomalous in not finding associations between methylmercury exposure and adverse effects. Second, both the National Research Council and the Environmental Protection Agency included the Seychelles Islands study in their analyses. The only way the acceptable level of methylmercury could be higher would be to ignore the two major positive studies that were modeled by the NRC, as well as six smaller studies, and rely solely on the single study showing no negative effects of methylmercury. Third, there is a substantial percentage of women of reproductive age in the United States with levels of methylmercury in their bodies above what EPA considers a safe level. As a result of this, over 300,000 newborns each year are exposed to methylmercury above levels U.S. EPA believes to be "safe". Fourth, increased exposure to methylmercury may result in atherosclerosis, heart attack, and even death from heart attack in men, suggesting that an additional large segment of the population may be at risk as a result of environmental methylmercury exposure.

Thank you for your time and attention.

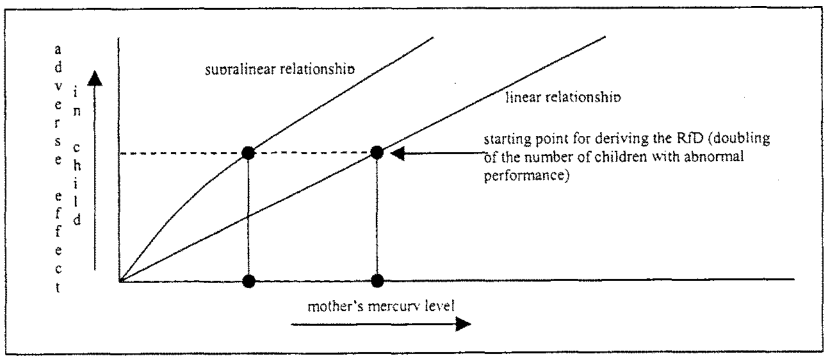
RESPONSES BY DEBORAH RICE TO ADDITIONAL QUESTIONS FROM SENATOR JEFFORDS

Question 1. In testimony, you indicated that "there might be virtually no safety factor at all" with respect to the effect level for mercury exposure. Does that mean that the reference dose should be lowered further? If so, what would be a safer and more protective reference dose?

Response. The current reference dose (RfD) is based on a cord blood mercury concentration associated with a defined risk: a doubling of the number of children performing in the abnormally low range. A total uncertainty factor of 10 was applied to account for inter-individual variability. There are several decisions made by EPA that, if different, would have resulted in a lower RfD.

(a) It was assumed that the ratio of cord-to-maternal blood mercury was one. Subsequent analyses of 10 studies revealed that cord blood has more mercury compared to maternal blood. The average ratio is 1.7:1.0, with the upper 5 percent of women having a ratio of 3.3:1.0. Based on just the average ratio, if no other decisions were changed, the RfD would be reduced from 1.0 μg/kg/day to 0.6 μg/kg/day.

(b) As was recommended by the NAS expert committee, EPA assumed that there was a linear relationship between adverse effects on a number of neuropsychological endpoints and the level of mercury in cord blood or maternal hair. In fact, the data from the Faroe Island study best fit a supra-linear model: i.e., the slope was actually greater at lower body burdens (see figure below). It turns out that this was also true for the New Zealand study. Recently, a study was published reporting a supra-linear shape to the relationship between adverse behavioral performance and blood lead levels in children. So this phenomenon, while somewhat counter-intuitive, may be real. Using the "best fit" model rather than forcing a linear relationship would result in a lower estimate of the defined adverse starting point (a doubling of the number of children performing in the abnormally low range), and thereby a lower RfD.



(c) EPA used a total uncertainty factor (UF) of 10 to derive the RfD, which is designed to provide a margin of safety against adverse effects. EPA typically applies an UF of 10 for inter-individual variability if the starting point is a no-observable-adverse-effect-level (NOAEL). If the starting point is the lowest level that has been demonstrated to produce an effect, with a NOAEL not identified, the EPA applies an additional UF, usually 10. In the case of methylmercury, even though the starting point was a level associated with an effect, only a total factor of 10 was applied, rather than the more typical 100. In addition, the UF of 10 for inter-individual variability is presumed to account for differences in both metabolism and response of the target organ (sensitivity) between individuals. The variability in metabolism of methylmercury between women has been demonstrated to be about 3. The variation in cord-maternal blood levels between women may be also about 3. These would be multiplied together to equal about 10. That allows no room for any variation in response of the fetal brain to methylmercury, which is undoubtedly not the case. Therefore a total UF of 10 is almost certainly inadequate to protect the most sensitive portion of the population.

The issue of whether the reference dose should be lowered, and if so, the appropriate value, requires thorough evaluation by a group of expert risk assessors and other scientists. Any new evaluation of the RfD should also include evaluation of the levels of methylmercury that produce adverse cardiovascular effects documented in several studies of adult males. It is currently unknown whether these effects occur at lower or higher levels than those that produce developmental neurotoxicity.

Question 2. What is a reasonable estimate of the approximate average mercury concentrations in non-commercial fish in the U.S.?

Response. EPA keeps an extensive data base of fish tissue contaminant levels from inland water bodies compiled by individual states (http://www.epa.gov/ost/fish/mercurydata.html). Data for average levels of mercury for 1987–2000 are in the attached figure. Average tissue levels vary significantly depending on species, such that deriving an "average" for all species is not particularly informative. Averages for different species range from 0.1 ppm for herring and whitefish to 0.9 ppm for bowfin. As can be seen from the figure, the average level for many species is below the 0.3 ppm level recommended by EPA (Water Quality Criterion for the Protection of Human Health: Methylmercury, OST, Office of Water, 2001, EPA–823-R–01–001). Approximately one third of species have average concentrations above this. Even for species with averages below 0.3 ppm, some samples will exceed this level. For species with averages about 0.5 ppm, more than half the samples will exceed the EPA recommended limit, whereas half the samples will exceed the 0.5 ppm action limit set by many European countries and Canada. Ocean fish and sharks can have levels that are considerably higher. For example, blue marlin average 3.08 ppm, with the highest level for an individual at 6.8 ppm (Florida Marine Research Institute Technical Reports' Mercury Levels in Marine and Estuarine Fishes of Florida. 1989–2001: FMRI Technical Report TR–9, Second Edition, Revised, 2003). Sharks such as white shark averaged over 5 ppm, with the highest value for a shark at 10 ppm (*ibid.*) These are non-commercial sport-caught species.

Question 3. You indicated that the NHANES data does not adequately capture the individuals or subpopulations that are likely to be the most exposed to non-commercial fish mercury concentrations above the reference dose. Are you aware of any work underway to collect this kind of data and hopefully protect these people from overexposure?

Response. There have been a number of relatively small studies focusing on fish intake by groups that consume large amounts of fish, specifically sports fishers and subsistence fishing communities. Most of these efforts have been by individual states or tribes. EPA is developing a data base of these studies, most of which are unpublished and not in the public domain, a project which I managed before leaving the agency. The data base currently includes about 70 studies (contact project officer Cheryl Itkin, EPA/ORD/National Center for Environmental Assessment, Washington, D.C. at itkin.cheryl@epa.gov).

There are also several published studies: Bellanger, T.M., Caesar, E.M., Trachtman, L. 2000. Blood mercury levels and fish consumption. J. La. Med. Soc, 152:64–73; Burge, P., Evans, S. 1994. Mercury contamination in Arkansas gamefish. A public health perspective. J. Ark. Med. Soc. 90:542–544; Hightower, J.M., Moore, D. 2003. Mercury levels in high-end consumers offish. Environ. Health Perspect. 111:604–608; and Knobeloch, L.M., Ziamik, M., Anderson, H.A., Dodson, V.N. 1995. Imported seabass as a source of mercury exposure: A Wisconsin case study. Environ. Health Perspect. 103:604–606.

Protecting individuals who may be at greater risk from over-exposure to methylmercury presents significant challenges. Forty states have fish advisories for inland waters, based largely on levels of mercury in fish. Some states have levels that are specific to particular water bodies, others have statewide advisories for all water bodies. Advisories typically are set with regard to species of fish, designating them as e.g. "no restriction", "eat no more than once a week", or "eat no more than once a month". If a person eats a fish from one restricted category they are meant not to eat fish from another restricted category in that month. Signs are posted by some states at specific water bodies, and most if not all states distribute literature related to fish advisories with fishing licenses. Some tribes have also performed significant outreach related to issues of contaminants in wild foods. Immigrant communities are often the most difficult to inform, as a result of language and cultural barriers. Minnesota, for example, has made a substantial effort to work with immigrant communities, publishing appropriate information in relevant languages, as well as performing extensive outreach activities. A few other states have made efforts in this regard as well. Some communities rely on fish as a significant protein source for both cultural and economic reasons. It is unfortunate indeed that these communities are risking adverse health outcomes by consuming what should be a very healthful food.

Question 4. Please describe the purposes and intended uses of the various Federal agencies' exposure limits for methyl mercury.

Response. EPA, FDA, and ATSDR have set exposure limits for methylmercury. The reference dose (RfD) set by EPA is designed to represent an "estimate of a daily exposure to the human population (including sensitive subgroups) that is likely to be without appreciable risk of deleterious [non-cancer] effects during a lifetime" (http://www.epa.gov/iris/index.html).

The minimal risk level (MRL) of ATSDR is "an estimate of the daily human exposure to a hazardous substance that is likely to be without appreciable risk of adverse noncancer health effects over a specified duration of exposure". MRLs may be derived for acute (1–14 days), intermediate (15–364 days) or chronic durations (over 364 days). ATSDR states that "[t]hese substance-specific estimates, which are intended to serve as screening levels, are used by ATSDR health assessors and other responders to identify health effects that may be of concern at hazardous waste sites. It is important to note that MRLs are not intended to define clean-up or action levels for ATSDR or other Agencies." [bold original] (http://www.atsdr.cdc.gov/mrls.html) It is critical to understand that ATSDR is involved in clean-up activities. The MRLs are designed to identify chemicals that are important for clean-up decisions. They are not intended as health-protective levels for the general population, or for a lifetime.

The FDA acceptable daily intakes (ADI) is "the amount of a substance that can be consumed daily over a long period of time without appreciable risk" (http://www.fda.gov; http://www.cfsan.fda.gov/-acrobat/hgstud16.pdf). For contaminants in food, FDA uses the ADI to derive an Action Level, "which defines the maximum allowable concentration of the contaminant in commercial food." In other words, the Action Level is supposed to be health-based.

The RfD and the ADI are designed to protect the general population from adverse effects from contaminants in food over a lifetime of exposure, including protection of sensitive populations. In contrast, the MRL is designed for a different purpose: identifying contaminants that may be important in making decisions regarding clean-up of contaminated sites.

The exposure limits from U.S. agencies are as follows:

EPA RfD: 0.1 μg/kg/day
ATSDR MRL: 0.3 μg/kg/day
FDA ADI: 0.4 μg/kg/day

Question 5. What is the preferred measurement methodology for most reliably determining and predicting the effect on children's developmental health of methyl mercury exposure?

Response. There has been considerable discussion within the academic and regulatory communities regarding what might be a "best" test or test battery for determining adverse neuropsychological function in children exposed to methylmercury. There are two basic strategies that have been used to assess methylmercury neurotoxicity. The first is the use of standard clinical instruments such as measures of IQ. These have the advantage of being standardized for the population, as well as assessing a wide range of functional domains. However, because they may be measuring a number of functions that are not affected in addition to those that are, the results can be "diluted", and therefore these tests may be less sensitive than a more focused approach. The second approach is to choose domain-specific tests based on the known effects of higher levels of the toxic chemical, if such effects are known. This strategy has the advantage of being potentially more sensitive than using broad-based clinical instruments. On the other hand, using domainspecific tasks runs the risk of looking at the wrong functions.

The investigators of the Faroe Islands study used a number of domain-specific tasks, based on the effects of high-level methylmercury exposure as well as the pathological changes in specific brain areas produced by methylmercury. The Faroe Island study found deficits in these' tasks. The investigators of the Seychelles study used standard clinical instruments that assessed a little bit of a lot of functions, which were standardized for a U.S. population rather than the Seychellois population. They found no effect of methylmercury. In contrast, the investigators of the New Zealand study, also using standard clinical instruments, did identify mercury-related deficits.

The consensus of the research community seems to be that a combination of both approaches should be used, The standard clinical instruments (e.g. full-scale IQ) are comprised of subscales (e.g. verbal, visuospatial) that can be used to explore more specific functional domains. Researchers should also use what is known about the behavioral and neuropathological effects of methylmercury to design domainspecific tests, with the hope that these will be maximally sensitive. To date, deficits in memory, language processing, visuospatial ability, motor function, and attention have been identified to be adversely affected by *in utero* methylmercury exposure. Hearing may also be adversely affected. New studies, or continued testing of current cohorts, should build on this knowledge to hone in even further on specific behavioral functions.

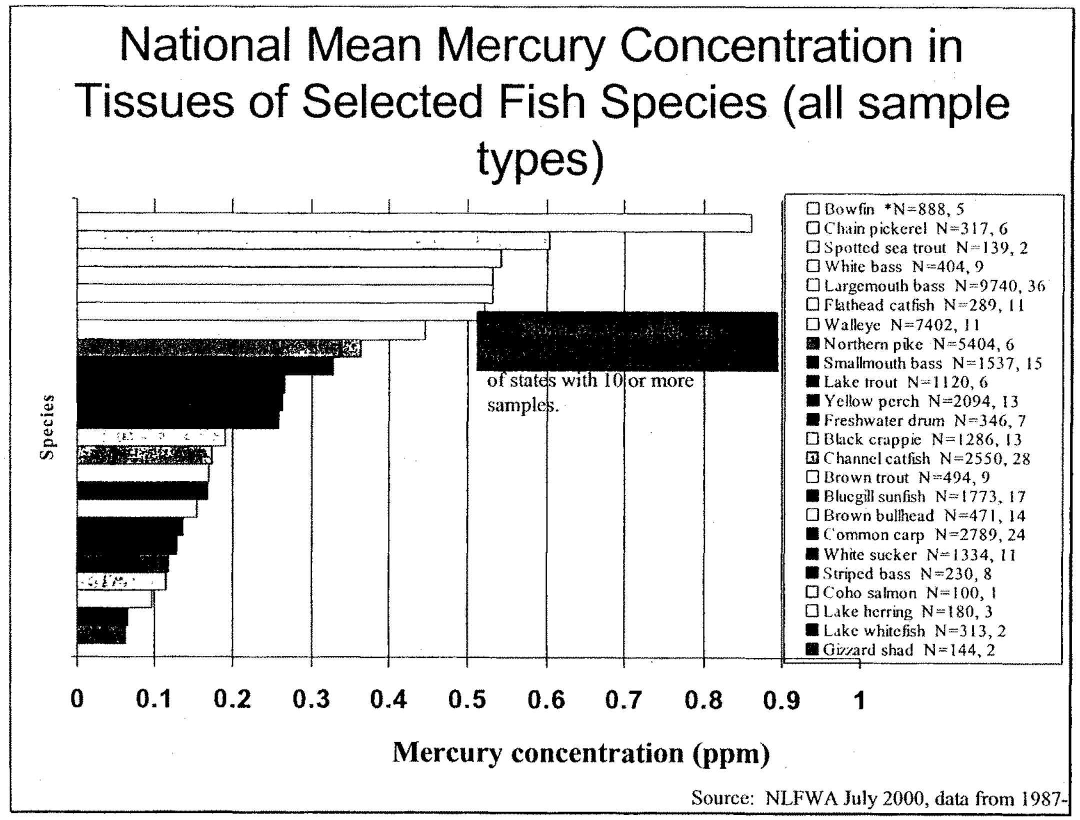
Question 6. In 1974, the FDA established a mercury action limit of .5 parts per million in fish. This was changed in 1979 to 1 part per million. What was the basis for this change?

Response. FDA set an action level of 0.5 ppm for mercury in fish in 1969,' in response to the recognition. of the devastating consequences of fetal exposure to methylmercury in the poisoning episodes in Minamata and Niigata, Japan. This level was reaffirmed in 1974, citing concerns about damage to the fetus at lower exposures than are harmful to the adult. The level was changed in 1979 as a result of a lawsuit by the fishing industry that resulted in a court ruling based on socioeconomic impacts presented by the National Marine Fisheries Service (NMFS). They argued that raising the action level would expand the number of fisheries available for exploitation and expand the profits of the fishing industry (Fed. Reg. 3990, 3992, 1979). The notice was a withdrawal of the proposed rulemaking and terminated a rulemaking procedure to codify the (then) existing action level limiting the amount of unavoidable mercury residues permitted in fish and shellfish of 0.5 ppm. The FR notice also indicates that "[t]he Food and Drug Administration will continue to monitor mercury levels in fish so that if there is any change in mercury residue levels as a result of raising the action level, or if there is any other change in the information regarding mercury in fish, the action level can be revised accordingly." Thus, the action limit is not health-based, but was established for economic considerations.

Question 7. What, if anything, should consumers of fish in the Great Lakes region and other areas that are downwind of major mercury emission sources such as coal-fired power plants, chlor-alkali manufacturing facilities and; waste incinerators, be advised to do with respect to limiting their methyl mercury exposure?

Response. Unfortunately, the majority of inland lakes and rivers are contaminated with mercury. Methylmercury is created from mercury by microorganisms in the water. Methylmercury is bioconcentrated as it is passed up the food chain, with older and larger fish at the top of the food chain containing more methylmercury than smaller fish or fish that are lower on the food chain. Methylmercury exposure in humans is exclusively from eating contaminated fish. Forty states have explicit fish advisories as a result of mercury contamination for consumption of fish based on species, size, and in some states specific water bodies. There were 2,242 advisories in 2000, up 8 percent o from 1999 and up 149 percent from 1993. By far the greatest number of fish advisories for mercury are around the Great Lakes and in the Northeastern states. Consumers are advised to carefully follow State fishing advisories for inland fish. There is an increasing recognition that commercial and/or ocean fish may represent a significant source of methylmercury exposure. Currently, FDA advises pregnant women, nursing mothers and young children against eating any shark, swordfish, tilefish, or king mackerel. Recent data indicate that canned white (albacore) tuna may have substantial levels of methylmercury, and so should be consumed seldom, especially by children. Other species such as fresh tuna and halibut may also have significant levels of methylmercury. Intake of purchased fish that are potentially high in methylmercury should be included by individuals in determining safe fish intake over a specific time period. In other words, consumers need to have detailed information on fish species from both commercial and non-commercial sources to keep track of their potential methylmercury intake.

This is an unsatisfactory solution, since fish should be a very healthful food. Moreover, sport fishing is an important economic resource in many areas, and some individuals rely on fishing for a substantial portion of their protein, particularly in certain immigrant communities. The ultimate solution is of course to decrease environmental deposition of mercury.


National Mean Mercury Concentration in Tissues of Selected Fish Species (all sample types)
Bowfin *N=888, 5
Chain pickerel N=317, 6
Spotted sea trout N=139, 2
White bass N=404, 9
Largemouth bass N=9740, 36
Flathead catfish N=289, 11
Walleye N=7402, 11
Northern pike N=5404, 6
Smallmouth bass N=1537, 15
Lake trout N=1120, 6
Yellow perch N=2094, 13
Freshwater drum N=346, 7
Black crappie N=1286, 13
Channel catfish N=2550, 28
Brown trout N=494, 9
Bluegill sunfish N=1773, 17
Brown bullhead N=471, 14
Common carp N=2789, 24
White sucker N=1334, 11
Striped bass N=230, 8
Coho salmon N=100, 1
Lake herring N=180, 3
Lake whitefish N=313, 2
Gizzard shad N=144, 2
[redacted]
of states with 10 or more samples.
Species
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Mercury concentration (ppm)
Source: NLFWA July 2000, data from 1987-

Risk Analysis, Vol. 23, No. 1, 2003

Methods and Rationale for Derivation of a Reference Dose for Methylmercury by the U.S. EPA

Deborah C. Rice,[1]* Rita Schoeny,[2] and Kate Mahaffey[3]

In 2001, the U.S. Environmental Protection Agency derived a reference dose (RfD) for methylmercury, which is a daily intake that is likely to be without appreciable risk of deleterious effects during a lifetime. This derivation used a series of benchmark dose (BMD) analyses provided by a National Research Council (NRC) panel convened to assess the health effects of methylmercury. Analyses were performed for a number of endpoints from three large longitudinal cohort studies of the neuropsychological consequences of *in utero* exposure to methylmercury: the Faroe Islands, Seychelles Islands, and New Zealand studies. Adverse effects were identified in the Faroe Islands and New Zealand studies, but not in the Seychelles Islands. The NRC also performed an integrative analysis of all three studies. The EPA applied a total uncertainty factor (UF) of 10 for intrahuman toxicokinetic and toxicodynamic variability and uncertainty. Dose conversion from cord blood mercury concentrations to maternal methylmercury intake was performed using a one-compartment model. Derivation of potential RfDs from a number of endpoints from the Faroe Islands study converged on 0.1 μg/kg/day, as did the integrative analysis of all three studies. EPA identified several areas for which further information or analyses is needed. Perhaps the most immediately relevant is the ratio of cord:maternal blood mercury concentration, as well as the variability around this ratio. EPA assumed in its dose conversion that the ratio was 1.0; however, available data suggest it is perhaps 1.5–2.0. Verification of a deviation from unity presumably would be translated directly into comparable reduction in the RfD. Other areas that EPA identified as significant areas requiring further attention are cardiovascular consequences of methylmercury exposure and delayed neurotoxicity during aging as a result of previous developmental or adult exposure.

KEY WORDS: Methylmercury; risk assessment; reference dose; Environmental Protection Agency; neuropsychological effects

Awareness of the neurotoxic potential of methylmercury resulted from a mass poisoning episode beginning in the 1950s in residents living near Minamata Bay in Japan. Methylmercury was discharged directly into the bay and was bioaccumulated and bioconcentrated by fish, which were a main dietary component of people in the region. Although both adult and infant or fetal exposure produced signs of methylmercury poisoning, the constellation of effects was different, and effects of developmental exposure were more severe.[1] Consequences of congenital methylmercury poisoning included mental retardation, cerebellar ataxia, seizures, visual abnormalities, and other neurological abnormalities. A

1 National Center for Environmental Assessment, Office of Research and Development, U.S. Environmental Protection Agency.
2 Office of Water, U.S. Environmental Protection Agency.
3 Office of Prevention, Pesticides and Toxic Substance, U.S. Environmental Protection Agency.
* Address correspondence to Deborah Rice, EPA/NCEA, MC 8623D, Washington, DC 20460; tel.: 202-564-3404; rice.deborah@epa.gov.

subsequent episode of methylmercury poisoning occurred in Niigata, Japan, in 1963–1965, involving hundreds of people; the source was a fertilizer factory that released methylmercury into a river that flows into a bay from which fish were caught.[2]

An episode of methylmercury poisoning occurred in Iraq in the 1970s when people consumed methylmercury-treated flour ground from grain intended for planting. Exposure was relatively acute compared to the poisoning episodes in Japan. Highly-exposed children manifested severe sensory impairment, paralysis, hyperactive reflexes, cerebral palsy, and impaired mental development.[3] Other children exposed *in utero* exhibited delayed walking and talking and an increased prevalence of neurological signs.[4] Subsequent epidemiological studies in various areas of the world including the Amazon,[5] Ecuador,[6] French Guiana,[7] Madeira,[8] and Canada (northern Quebec)[9] documented adverse neuropsychological effects associated with developmental exposure to methylmercury, including sensory, motor, and/or cognitive deficits.

The U.S. Environmental Protection Agency (EPA) has, as part of its mandate, a responsibility to perform risk assessments for chemicals present in the environment that may pose a hazard to human health. Risk assessment may be defined as "the characterization of the potential adverse health effects of human exposure to environmental hazards."[10] To accomplish this for noncancer effects, a reference dose (RfD) is derived, defined as "an estimate (with uncertainty spanning perhaps an order of magnitude) of a daily oral exposure to the human population (including sensitive subgroups) that is likely to be without appreciable risk of deleterious effects during a lifetime." The RfD may be derived from a no adverse effect level (NOAEL) (the highest dose at which no adverse effects are identified), or a low adverse effect level (LOAEL) (the lowest dose at which adverse effects are observed), typically identified in animal studies. It may also be derived from benchmark dose (BMD) analysis, which was used by EPA for derivation of the RfD for methylmercury described herein. In any case, a point of departure (POD) is identified, which is the point on the dose-response curve that marks the starting point for derivation of the RfD. Typically, this POD is divided by one or more uncertainty factors (UFs) in order to account for various extrapolations and perceived insufficiencies in the database.[11]

EPA derived an RfD for methylmercury in 1995 based on the episode of methylmercury poisoning in Iraq. An exposure-effect curve was modeled based on a composite of adverse effects in 81 infants, including onset of walking and talking, soft neurological signs, mental symptoms, and seizures, using a Weibull model. The RfD was 0.1 μg/kg/day of methylmercury, based on estimated intake by the mothers during pregnancy. In 1997, Congress mandated that EPA fund an expert panel under the auspices of the National Research Council (NRC) to determine whether the RfD was scientifically justifiable. The NRC panel concluded that an RfD of 0.1 μg/kg/day was scientifically justifiable based on its analysis.[12] However, NRC recommended that EPA not base its assessment on the Iraqi study, but rather on new large epidemiological studies that had become available since 1995. EPA therefore completed a risk assessment of methylmercury in 2001, using the studies and the analyses suggested by the NRC.

1. STUDIES CONSIDERED IN DERIVATION OF THE RfD

There are three recent epidemiological studies suitable for quantitative analysis that have become available in the peer-reviewed literature since EPA's derivation of an RfD in 1995, and that were used in the extensive quantitative analyses by the NRC.[12] These longitudinal prospective developmental studies were conducted in the Seychelles Islands, the Faroe Islands, and New Zealand. The Seychelles Islands study consisted of 779 mother-infant pairs from a fish-eating population.[13–18] Infants were followed from birth to 5.5 years of age, and assessed at various ages on a number of standardized neuropsychological endpoints. The independent variable was maternal hair mercury concentrations. The Faroe Islands study included about 900 mother-infant pairs; children were tested on a variety of tasks at 7 years of age.[19] The main independent variable was cord blood mercury, although maternal hair mercury was also measured. In the New Zealand study, 38 children of mothers with hair mercury levels during pregnancy greater than 6 ppm were matched with children whose mothers had lower hair mercury concentrations.[20,21] At 6 years of age, a total of 237 children were assessed on a number of neuropsychological endpoints similar to those used in the Seychelles study.[20] Investigators in the Seychelles Islands study reported no evidence of impairment related to *in utero* methylmercury exposure in their main study, whereas the other two studies

Table I. Tests Modeled by NRC, Functions Assessed, and Potential Societal Relevance

Study	Test	Domain/Function Assessed	Societal Relevance
Seychelles	Bender Copying Errors	Visuospatial	Math performance
	McCarthy GCI	Full-Scale IQ	School performance, intelligence
	WJ Applied Problems	Ability to solve problems	Academic skills
	CBCL	Social and adaptive behavior	Antisocial behavior, need for therapeutic services
	Preschool Language Scale	Broad-based language	Learning, intelligence, school performance
	WJ Letter/Word Recognition	Word recognition	Reading ability, school performance
Faroes	Finger Tapping	Motor performance	Motor speed/neuropathy
	CPT Reaction Time	Vigilance, attention, information processing speed	Intelligence, school behavior, and performance
	Bender Copying Errors	Visuospatial	Math performance
	Boston Naming Test	Expressive vocabulary	Reading, school performance
	CVLT: Delayed Recall	Memory	Learning ability, school performance
New Zealand	TOLD Language Development	Broad-based language	Literacy skills, learning, school performance
	WISC-R: PIQ	Performance IQ, e.g., visuospatial, sustained attention, sequential memory	Learning, school performance
	WISC-R: FSIQ	Full-scale IQ, e.g., PIQ + verbal processing, expressive vocabulary	Learning, school performance
	McCarthy Perceptual Performance	Performance IQ, e.g., visuospatial, audition, memory	Learning, school performance
	McCarthy Motor Test	Gross and fine motor skills	Motor system integration

Notes: WJ = Woodcock-Johnson Tests of Achievement; CBCL = Child Behavior Check List; CPT = Continuous Performance Test; CVLT = California Verbal Learning Test; TOLD = Test of Language Development; WISC-R:PIQ = Wechsler Intelligence Scale for Children-Revised Performance IQ; WISC-R FSIQ = Wechsler Intelligence Scale for Children-Revised Full-Scale IQ.
Source: Reference 26.4–51.

found exposure-related effects on a number of neuropsychological endpoints.

The NRC quantitative analysis included five endpoints from the Faroe Islands study from a total of nine that had been reported as significantly affected by methylmercury exposure[19] (Table I). Similarly, five endpoints negatively associated with methylmercury exposure in the New Zealand study[20] were used in the quantitative analysis by the NRC. The study in the Seychelles Islands did not identify adverse effects of methylmercury exposure on any of the neuropsychological endpoints; the NRC modeled all six of the endpoints assessed in that study. The tests used in the Faroe Islands study were chosen to assess specific functional domains. In contrast, the tests used in the New Zealand and Seychelles Islands studies were apical tests that yield global scores representing the integration of performance over a number of domains. A workgroup convened by NIEHS in 1998 to compare the Faroe and Seychelles Islands studies postulated that differences in the test endpoints assessed might account for the conflicting findings between the two studies.[22] However, the New Zealand study also used apical tests, and the NRC panel concluded that the difference in the results between the Seychelles and Faroe Islands studies was probably not due to differences in assessment methodology.[12] All these endpoints assessed functions that are important for the child's ability to learn, remember, and be successful in an academic setting.

Maternal exposure to methylmercury in the studies was comparable and is unlikely to account for the differential findings in the studies. The maternal hair geometric mean in the Faroe Islands study was 4.3 ppm, with an interquartile range of 2.6–7.7 ppm. The mean maternal hair mercury in the Seychelles Islands study was 6.8 ppm with a standard deviation of 4.5. In the New Zealand study, the mean hair mercury concentration was 8.3 ppm for the 38 mothers in the "high" group, each matched to three "controls" with lower mercury levels. There are other differences between the studies that may account for the

differences in results, which have been discussed extensively elsewhere.[12,22] Among these are children's age at testing, differential genetic susceptibility of the populations, differential pattern of exposure to methylmercury (episodic vs. relatively continuous), and co-exposure to PCBs in the Faroe Islands study (discussed below). It is not possible to determine on the current evidence which if any of these factors is responsible for the different findings in the Seychelles versus Faroe Islands and New Zealand studies.[12,22] NRC[12] points out that the power of the Seychelles study to detect the small effects identified in the Faroe Islands study was only about 50%.

2. DETERMINATION OF THE POINTS OF DEPARTURE FOR THE RfD

EPA chose benchmark dose (BMD) analysis as the most appropriate method of quantifying the dose-effect relationship in these studies, which was the recommendation of the NRC.[12] The K-power model was used to model the exposure-effect functions. This model allows for a decreased magnitude of response at lower exposures compared to higher exposures (sublinearity), which would be appropriate if there were a threshold, for example. The K-power model was constrained to $K \geq 1$, which precluded a greater apparent response at low exposures relative to higher exposures (supralinearity). $K = 1$ (i.e., a linear dose-effect function) provided the best fit of the K-power models for the Faroe Islands data;[23,24] therefore this model was used in the NRC BMD calculations for all three studies.

Benchmark dose analysis requires two additional decisions once an appropriate model has been chosen. When continuous data are used, a point on the curve below which responses are considered "abnormal" must be chosen, termed P_0. A value of $P_0 = 0.05$ was used in the EPA assessment; that is, the cutoff for abnormal response was set at the lowest 5% (5th percentile) of children. Most human characteristics, including children's neurodevelopmental abilities, have an approximately bell-shaped or "normal" distribution. Generally speaking, children who function at or below approximately the 5th percentile would be considered significantly developmentally compromised for the ability that is being measured (e.g., an IQ < 75, or performance on standard scores for more specific abilities such as attention, language, or memory).

The second decision that must be made is the choice of the increase in the proportion of individuals that will be expected to perform in the "abnormal" category in an exposed versus an unexposed population. This is defined as the benchmark response (BMR). A BMR of 0.05 was chosen for this assessment, which would result in a doubling of the number of children with a response at or below the 5th percentile in an unexposed population. Effects were identified on a number of neuropsychological endpoints at approximately the same body burden, which included tests that are predictive of reading and mathematics performance, overall academic performance, and antisocial behavior. The NAS concluded that "[d]eficits of the magnitude reported in the [Faroe and New Zealand] studies are likely to be associated with increases in the number of children who have to struggle to keep up in a normal classroom or who might require remedial classes or special education."[12:325] Such adverse effects are costly for the individual as well as for society as a whole. The choice of a BMR greater than 0.05, which potentially would allow a greater proportion of children to perform in the "abnormal" range, was considered unjustifiable by EPA.

BMDs were calculated for each of the endpoints described above for each of the three studies (Table II). The lower limit on the 95% confidence interval of the BMD (the BMDL) was calculated for each endpoint. These BMDLs served as potential points of departure for the RfD. The BMDLs from the Faroe Islands study were 12–15 ppm in maternal hair, whereas those in the New Zealand study were 4–6 ppm. BMDLs from the Seychelles Islands study were 17–25, about 50% higher than those in the Faroe Islands and 250–300% higher than those from the New Zealand study. It is interesting that the most sensitive measure in the Seychelles Islands study was the Child Behavior Checklist, which is a test of social and adaptive abilities.

It is important to recognize that values derived from a BMD analysis do not represent a threshold, nor are they comparable to a NOAEL or LOAEL as typically derived from animal studies. NOAELs and LOAELs are dependent on the doses chosen for the study. BMD analysis, on the other hand, makes full use of the range of the exposure-effect curve. BMD analysis identifies body burdens that are associated with an identifiable increased risk of a specific adverse effect. Moreover, in the Faroe Islands study at least, there was no evidence for a threshold for adverse effects within the range of maternal exposures in the study, as evidenced by the fact that the $K = 1$ model provided a better fit to the data than K values greater than 1. In

Table II. BMD (BMDL) Estimates (ppm MeHg in Maternal Hair) for Endpoints from Three Longitudinal Prospective Studies

Study	Endpoint	BMD[a]	BMDL
Seychelles	Bender Copying Errors	***	25
	Child Behavior Checklist	21	17
	McCarthy General Cognitive	***	23
	Preschool Language Scale	***	23
	WJ Applied Problems	***	22
	WJ Letter/Word Recognition	***	22
Faroe Islands	Finger Tapping	20	12
	CPT Reaction Time	17	10
	Bender Copying Errors	28	15
	Boston Naming Test	15	10
	CVLT: Delayed Recall	27	14
New Zealand	TOLD Language Development	12	6
	WISC-R:PIQ	12	6
	WISC-R:FSIQ	13	6
	McCarthy Perceptual Performance	8	4
	McCarthy Motor Test	13	6

[a]BMDs are calculated from the K-power model under the assumption that 5% of the responses will be abnormal in unexposed subjects ($P_0 = 0.05$), assuming a doubling of the excess risk (BMR = 0.05).
Notes: WJ = Woodcock-Johnson Tests of Achievement; CPT = Continuous Performance Test; CVLT = California Verbal Learning Test; TOLD = Test of Language Development; WISC-R:PIQ = Wechsler Intelligence Scale for Children-Revised Performance IQ; WISC-R:FSIQ = Wechsler Intelligence Scale for Children-Revised Full-Scale IQ.
Source: Reference 12:284.

addition, the logarithmic model provided a better fit than the $K = 1$ model for at least some endpoints in the Faroe Islands study.[23,24] In all cases, the logarithmic model yielded BMDLs considerably lower than the $K = 1$ model. EPA used the $K = 1$ model because of concerns about the stability of the logarithmic model at low exposures, as well as questions concerning the biologic plausibility of a supralinear response at low mercury exposures. However, these analyses provide further evidence against a threshold within the range of exposures in the Faroe Islands study.

In addition to methylmercury, the Faroe Islands cohort was also exposed to polychlorinated biphenyls (PCBs). An association was identified between PCB levels in cord tissue and four of the nine measures for which an effect of methylmercury was also identified.[19] Statistical analyses performed by the Faroe investigators suggested that the effects of methylmercury and PCBs were independent of each other.[25] NRC presented additional analyses designed to explore the influence of PCBs (Table III). PCB concentrations were determined on cord tissue from about one-half of the Faroese cohort (about 450 children) and the "adjusted for PCBs" values are based on those children for whom PCB cord tissue levels and cord blood mercury concentrations were available. Analyses were also performed on children in the lowest tertile with respect to PCB levels, which reduced the number of children available for analyses to about 150. No pattern was apparent for the PCB-adjusted analyses compared to the original results, and the observed variability was probably no more than that which would be expected by chance alone. This analysis provides compelling evidence that the effects of methylmercury identified in this study were not the consequence of exposure to PCBs.

In addition to determining the BMDLs for individual endpoints in each of the three studies, NRC also used a hierarchical random effect model to reduce random variation in the estimate for these same endpoints from all three studies.[12:290–294, Table 7-5] Additionally, this analysis was used in calculating BMD and BMDLs for the most sensitive and median endpoints from both the Faroe Islands and New Zealand studies.[12:294, Table 7-6] This approach also allowed an integrative analysis based on all three studies.

3. DERIVATION OF THE RfD

To derive the RfD, the level of mercury in cord blood had to be converted to a corresponding intake of methylmercury by the mother. EPA used a one-compartment model for maternal methylmercury kinetics in the assessment.[26] This model provided a reasonable fit to the data,[27] and did not require a series of assumptions necessary for other models.[28] The BMDLs from the endpoints from all three studies modeled by the NRC were converted using a one-compartment model to an ingested dose of methylmercury associated with a corresponding cord blood level (Table IV).

The last step in the derivation of an RfD is the choice of UFs. Since the methylmercury assessment was based on BMD analysis of developmental effects in humans, some of the UFs often employed were irrelevant (e.g., extrapolation from animals to humans, from less-than-chronic to chronic, from LOAEL to NOAEL). The assessment utilized a rich database for developmental neurotoxicity, such that EPA believed that an UF for insufficiency in the database was unwarranted. However, areas where additional

Exposure	Endpoint	Full Cohort	Adjusted for PCBs	Low-PCB Tertile
Maternal hair (ppm)	Finger Tapping	20 (12)	17 (9)	7 (4)
	CPT Reaction Time	18 (10)	27 (11)	13 (5)
	CVLT: Delayed Recall	27 (14)	39 (12)	32 (7)
Cord Blood (ppb)	Finger Tapping	140 (79)	149 (66)	41 (24)
	CPT Reaction Time	72 (46)	83 (49)	53 (28)
	BNT	85 (58)	184 (71)	127 (40)
	CVLT: Delayed Recall	246 (103)	224 (78)	393 (52)

[a]BMDs are calculated under the assumption that 5% of the responses will be abnormal in unexposed subjects ($P_0 = 0.05$), assuming a doubling of the excess risk (BMR = 0.05).
Notes: CPT = Continuous Performance Test; BNT = Boston Naming Test; CVLT = California Verbal Learning Test.
Source: Reference 12:289, Table 7-4.

Table III. BMD (BMDL) Estimates from the Faroe Islands Study With and Without Adjustment for PCBs and in the Subset of Children in the Lowest Tertile with Respect to PCB Exposure (Calculated Using the *K*-Power Model)

information or analyses would be valuable were identified (see below). The only UF factor included in the assessment was a factor of 10 for intrahuman variability.

A composite UF of 10 was used by EPA, $10^{0.5}$ for pharmacokinetic (absorption, elimination, and tissue distribution) variability and uncertainty, and $10^{0.5}$ for pharmacodynamic (biologic response of the relevant organ or tissue) variability and uncertainty. (Half-log units are expressed as 3.0 rather than $10^{0.5}$ by EPA.) For the former, EPA relied on the NRC analyses of variability in the pharmacokinetic factors underlying the conversion from a maternal body burden of methylmercury to an ingested daily dose of methylmercury that corresponded to that concentration.[12:92, Table 3-1] The interindividual variability in ingested dose corresponding to a given maternal body burden from three independent analyses was 1.7–3.0 based on maternal blood, and 1.8–3.3 for maternal hair, using the ratio of the 50th percentile/1st percentile. EPA therefore considered that a factor of three was protective of 99% of the population based on variability of maternal blood mercury half-life.

A quantitative uncertainty analysis was not possible for toxicodynamics. However, the population of the Faroe Islands is descended from Scandinavian stock that settled many generations ago and is extremely homogeneous. The population of the Seychelles Islands is a stable European-African population that may also be more homogeneous than that of the United States. The New Zealand cohort, on the other hand, included individuals from several ethnic groups, and BMDs calculated from this population were lower than those of the Faroe or Seychelles Islands. The relative variability within and between these different populations is unknown, as is the comparability of these populations to that of the United States. A three-fold UF for toxicodynamic variability and uncertainty was applied by EPA.

Support for the choice of an overall UF of 10 is provided by the additional analyses of the Faroese neuropsychological data.[19] Associations remained significant when the observations from offspring of mothers with maternal hair mercury concentrations greater than 10 ppm were excluded from the analyses. This finding indicates that methylmercury-related deficits are present at concentrations below 10 ppm mercury in hair, which corresponds to about 40 ppb in cord blood. The points of departure for the Faroe Islands study were 12–15 ppm in maternal hair, based on the BMD analysis. It may be that effects below 10 ppm are less severe than the magnitude defined by the choices of P_0 and BMR in the BMD analysis. In any event, the fact that the associations were still identified at hair levels below 10 ppm strongly indicates the appropriateness of an UF of at least 10.

The last column of Table IV shows the RfDs for the various endpoints including application of an UF of 10. The calculated RfD values converge on the same point: 0.1 μg/kg/day. Among the endpoints from the Faroe Island study, there are three deviations from 0.1 μg/kg/day: 0.2 μg/kg/day for the CVLT, entire cohort; 0.05 μg/kg/day for CPT, lowest PCB tertile; and 0.05 μg/kg/day for Finger Tapping, lowest PCB tertile. EPA also calculated geometric means from the four endpoints from the Faroe Islands study;[26] RfDs are 0.1 μg/kg/day based on these calculations. For the New Zealand study, smoothed values, both the median value and the results of the McCarthy Perceived

Table IV. BMDLs, Ingested Dose, and RfDs for Various Endpoints from the Faroes Islands, New Zealand, and the NRC Integrative Analysis[a]

Test[b]	BMDL ppb Mercury Cord Blood	Ingested Dose μg/kg/day[c]	RfD μg/kg/day[d]
BNT Faroes			
Whole cohort	58	1.081	0.1
PCB adjusted	71	1.323	0.1
Lowest PCB	40	0.745	0.1
CPT Faroes			
Whole cohort	46	0.857	0.1
PCB adjusted	49	0.913	0.1
Lowest PCB	28	0.522	0.05
CVLT Faroes			
Whole cohort	103	1.920	0.2
PCB adjusted	78	1.454	0.1
Lowest PCB	52	0.969	0.1
Finger Tap Faroes			
Whole cohort	79	1.472	0.1
PCB adjusted	66	1.230	0.1
Lowest PCB	24	0.447	0.05
Geometric mean Faroes			
Whole cohort	68	1.268	0.1
PCB adjusted	65	1.212	0.1
Lowest PCB	34	0.634	0.1
Smoothed values			
BNT Faroes	48	0.895	0.1
CPT Faroes	48	0.895	0.1
CVLT Faroes	60	1.118	0.1
Finger Tap Faroes	52	0.969	0.1
MCCPP New Zealand	28	0.522	0.05
MCMT New Zealand	32	0.596	0.1
Median values			
Faroes	48	0.895	0.1
New Zealand	24	0.447	0.05
Integrative			
All endpoints	32	0.596	0.1

[a] $BMDL_{05}$s from Reference 12:Tables 7-4, 7-5, 7-6. Hair mercury was converted to blood mercury using a 250:1 ratio and an assumption of equivalent maternal and cord levels.
[b] BNT = Boston Naming Test; CPT = Continuous Performance Test; CVLT = California Verbal Learning Test; MCCPP = McCarthy Perceived Performance; MCMT = McCarthy Motor Test.
[c] Calculated using a one-compartment model.
[d] Calculated using an UF of 10.
Source: Reference 26:4–61.

Performance test, yield RfDs of 0.05 μg/kg/day, and the McCarthy Motor Test yields an RfD of 0.1 μg/kg/day. Based on the integrative analysis of all three studies, the RfD would be 0.1 μg/kg/day.

Rather than choose a single measure for the RfD critical endpoint, EPA based the derived RfD on the BMDLs for a number of endpoints from the Faroe Islands study, with supporting analyses from the New Zealand study and the integrative analysis of all three recent large epidemiological studies.

4. OTHER ISSUES NOT ADDRESSED IN THE RfD

In addition to the issues considered to be included in the overall uncertainty factor of 10, EPA recognized additional areas for which data are lacking, and for which there is a reasonable possibility that further information could be important in the RfD determination. A potentially very important issue is the ratio of maternal blood mercury concentrations to the concentrations in fetal cord blood. EPA assumed that the ratio was 1:1 in its dose conversion from cord blood mercury concentrations to methylmercury daily intake by the mother. However, there is evidence that the cord:maternal blood ratio is greater than 1. Review of the literature identified about 20 studies that reported cord and maternal blood mercury concentrations.[26] The composite ratio from the studies indicates that the cord:maternal blood ratio was around 1.7. The calculated ratios are means and do not reflect the full range of variability in the individual mother-fetal pairs. For example, Vahter *et al.*[29] reported the 5th and 95th percentiles of cord:maternal Hg to be 0.88 and 3.1. A deviation from EPA's assumption that cord:maternal blood mercury ratio was 1 would be directly reflected in inaccuracy of the RfD. For example, if the real ratio were in fact 1.7, the RfD should be decreased by a directly comparable amount (to 0.06 μg/kg/day). This issue requires further investigation to determine both the best estimate of central tendency (average) as well as the variability in the ratio for individual mother:infant pairs. The latter calculation would have relevance for the choice of the uncertainty factor: specifically, the factor used for pharmacokinetic variability and uncertainty.

Another area that EPA identified for further analysis was the association between methylmercury exposure and adverse cardiovascular effects. In a study of 1,000 seven-year-old Faroese children, diastolic and systolic blood pressures increased by 13.9 and 14.6 mm Hg, respectively, as the cord-blood mercury increased from 1 to 10 μg/L.[30] A 47% decrease in heart rate variability (an indication of decreased cardiac autonomic control) was also observed. In a seven-year observation period of 1,833 Finnish men, individuals with hair mercury in the highest tertile (2 ppm or higher) had a 2.0 times greater risk of acute

myocardial infarction compared to the rest of the study population.[31] The body burden at which cardiovascular effects are observed in adults should be identified, preferably with quantitative procedures such as BMD analysis.

An additional area of concern for which quantitative data are lacking is the onset or exacerbation of neurological deficits in aging populations previously exposed to methylmercury. "Acts of daily living," which included the abilities to independently eat, bathe, wash, dress, and use the toilet, were evaluated in people diagnosed with Minamata Disease (MD).[32] The prevalence of deficits was relatively greater in persons with MD compared with controls as a function of increasing age. In other words, exposure to methylmercury three decades earlier accelerated the aging process in aged individuals relative to younger ones. Adults who lived in a methylmercury-polluted area near Minamata City in Kumamoto Prefecture in Japan but who were not diagnosed with MD were compared with age-matched adults from a region with low exposure.[33] Complaints that were significantly higher in methylmercury-contaminated areas included a number of neurological signs and symptoms. Animal studies lend support to the conclusion that methylmercury can have delayed effects that are uncovered with age. Neurological and immunological impairment emerged during middle adulthood in mice exposed prenatally to methylmercury.[34] Rats exposed to methylmercury *in utero* and post-natally exhibited a decline in performance in a task that required a substantial motor output at an earlier age than did control rats.[35] Rice and colleagues[36–38] identified accelerated aging of sensory system function in a series of studies in monkeys exposed developmentally to methylmercury. All these observations suggest that either developmental or adult exposure to methylmercury can have adverse long-term sequelae that may not be detected for years or decades following cessation of exposure. However, these effects cannot be quantified in humans based on available data.

5. CONCLUSION

The 2001 EPA RfD for methylmercury is 0.1 μg/kg/day, based on neuropsychological deficits resulting from *in utero* exposure in humans. The RfD is derived from BMD analyses of a number of endpoints from a large longitudinal cohort study in the Faroe Islands, and an integrative analysis of three longitudinal studies. Analyses of a longitudinal study in New Zealand provide supporting evidence that the derived RfD is appropriate. EPA considers that the RfD is for the entire population, not only for women of childbearing age. It is clear that infants and children are more sensitive than adults to methylmercury-induced neurotoxicity. In addition, the potential for cardiovascular effects and delayed neurotoxicity in adults, which are at present insufficiently quantified, suggests that adherence to the RfD by everyone is the most health-protective strategy.

ACKNOWLEDGMENTS

This article was written solely by the senior author. The document for the derivation of the reference dose for methylmercury, written by the authors, is Chapter 4 of the *Water Quality Criterion for the Protection of Human Health: Methylmercury*, EPA-823-R-01-001, 2001. A summary of the reference dose derivation, written by the senior author, is on the EPA Integrated Risk Information System (IRIS) website. The opinions expressed in this manuscript do not necessarily represent official EPA policy.

REFERENCES

1. Harada, M. (1995). Minamata disease: Methylmercury poisoning in Japan caused by environmental pollution. *Critical Reviews in Toxicology*, *25*(1), 1–24.
2. Tsubaki, T. (1972). Follow-up studies of Niigata-Minamata Disease. *Kagaku*, *42*, 526–531.
3. Amin-Zaki, L., Elhassani, S., Majeed, M. A., Clarkson, T. W., Doherty, R. A., & Greenwood, M. (1974). Intra-uterine methylmercury poisoning in Iraq. *Pediatrics*, *54*, 587–595.
4. Marsh, D. O., Clarkson, T. W., Cox, C., *et al.* (1987). Fetal methylmercury poisoning: Relationship between concentration in single strands of maternal hair and child effects. *Archives of Neurology*, *44*, 1017–1022.
5. Grandjean, P., White, R. F., Nielsen, A., Cleary, D., & de Oliveira Santos, E. (1999). Methylmercury neurotoxicity in Amazonian children downstream from gold mining. *Environmental Health Perspectives*, *107*, 587–591.
6. Counter, S. A., Buchanan, L. H., Laurell, G., & Ortega, F. (1998). Blood mercury and auditory neuro-sensory responses in children and adults in the Nambija gold mining area of Ecuador. *Neurotoxicology*, *19*(2), 185–196.
7. Cordier, S., Garel, M., Mandereau, L., Marcel, H., Doineau, P., Gasme-Seguret, S., Josse, P., White, R., & Amiel-Tison, C. (2002). Neurodevelopmental investigations among methymercurey-exposed children in French Guiana. *Environmental Research*, Section A 89, 1–11.
8. Murata, K., Weihe, P., Renzoni, A., Debes, F., Vasconcelos, R., Zino, F., Araki, S., Jorgensen, P. J., White, R. F., & Grandjean, P. (1999). Delayed evoked potentials in children exposed to methylmercury from seafood. *Neurotoxicology and Teratology*, *21*(4), 343–348.
9. McKeown-Eyssen, G., & Ruedy, J. (1983). Prevalence of neurologic abnormality in Cree Indians exposed to methylmercury in northern Quebec. *Clinical and Investigative Medicine*, *6*, 161–169.

10. National Research Council (NRC). (1983) *Risk assessment in the federal government: Managing the process*. Washington, DC: NAS Press.
11. U.S. EPA. (1987). *The risk assessment guidelines of 1986*. EPA/600/8-87/045. Washington, DC: Office of Health and Environmental Assessment.
12. National Research Council (NRC). (2000). *Toxicological effects of methylmercury*. Committee on the Toxicological Effects of Methylmercury, Board on Environmental Studies and Toxicology, Commission on Life Sciences, National Research Council. Washington, DC: National Academy Press.
13. Myers, G. J., Davidson, P. W., Shamlaye, C. F., Axtell, C. D., Cernichiari, E., Choisy, O., Choi, A., Cox, C., & Clarkson, T. W. (1997). Effects of prenatal methylmercury exposure from a high fish diet on developmental milestones in the Seychelles Child Development Study. *Neurotoxicology*, *18*(3), 819–830.
14. Myers, G. J., Marsh, D. O., Cox, C., Davidson, P. W., Shamlaye, C. F., Tanner, M. A., Choi, A., Cernichiari, E., Choisy, O., & Clarkson, T. W. (1995). A pilot neurodevelopmental study of Seychellois children following *in utero* exposure to methylmercury from a maternal fish diet. *Neurotoxicology*, *16*(4), 629–638.
15. Myers, G. J., Marsh, D. O., Davidson, P. W., Cox, C., Shamlaye, C. F., Tanner, M. A., Choi, A., Cernichiari, E., Choisy, O., & Clarkson, T. W. (1995). Main neurodevelopmental study of Seychellois children following *in utero* exposure to methylmercury from a maternal fish diet: Outcome at six months. *Neurotoxicology*, *16*(4), 653–664.
16. Myers, G. J., Davidson, P. W., Cox, C., Shamlaye, C. F., Tanner, M. A., Choisy, O., Sloane-Reeves, J., Marsh, D. O., Cernichiari, E., Choi, A., Berlin, M., & Clarkson, T. W. (1995). Neurodevelopmental outcomes of Seychellois children sixty-six months after *in utero* exposure to methylmercury from a maternal fish diet: Pilot study. *Neurotoxicology*, *16*(4), 639–652.
17. Davidson, P. W., Myers, G. J., Cox, C., Axtell, C., Shamlaye, C., Sloane-Reeves, J., Cernichiari, E., Needham, L., Choi, A., Yang, Y., Berlin, M., & Clarkson, T. W. (1998). Effects of prenatal and postnatal methylmercury exposure from fish consumption on neurodevelopment: Outcomes at 66 months of age in the Seychelles child development study. *Journal of The American Medical Association*, *280*, 701–707.
18. Davidson, P., Myers, G., Cox, C., Shamlaye, C., Marsh, D., Tanner, M., Berlin, M., Sloane-Reeves, J., Cernichiari, E., Choisy, O., Choi, A., & Clarkson, T. (1995). Longitudinal neurodevelopmental study of Seychellois children following in utero exposure to methylmercury from maternal fish ingestion: Outcomes at 19 and 29 months. *Neurotoxicology*, *16*, 677–688.
19. Grandjean, P., Weihe, P., White, R., Debes, F., Arak, S., Yokoyama, K., Murata, K., Sorensen, N., Dahl, R., & Jorgensen, P. (1997). Cognitive deficit in 7-year-old children with prenatal exposure to methylmercury. *Neurotoxicology and Teratology*, *20*, 1–12.
20. Kjellstrom, T., Kennedy, P., Wallis, S., & Mantell, C. (1989). *Physical and mental development of children with prenatal exposure to mercury from fish. Stage 2: Interviews and psychological tests at age 6*. Report 3642. Solna, Sweden: National Swedish Environmental Protection Board.
21. Kjellstrom, T., Kennedy, P., Wallis, S., & Mantell, C. (1986). *Physical and mental development of children with prenatal exposure to mercury from fish. Stage 1: Preliminary test at age 4*. Report 3080. Solna, Sweden: National Swedish Environmental Protection Board.
22. National Institute of Environmental Health Sciences (NIEHS). (1999). Scientific issues relevant to assessment of health effects from exposure to methylmercury. Workshop organized by Committee on Environmental and Natural Resources (CENR) Office of Science and Technology Policy (OSTP), The White House. Raleigh, NC.
23. Budtz-Jørgensen, E., Grandjean, P., Keiding, N., *et al.* (2000). Benchmark dose calculations of methylmercury-associated neurobehavioral deficits. *Toxicol. Lett.*, *112–113*, 193–199.
24. Budtz-Jørgensen, E., Keiding, N., & Grandjean, P. (1999). *Benchmark modeling of the Faroese methylmercury data*. Final Report to U.S. EPA. Research Report 99/5. Copenhagen: Department of Biostatistics, University of Copenhagen
25. Budtz-Jørgensen, E., Keiding, N., Grandjean, P., & White, R. (1999). Methylmercury neurotoxicity independent of PCB exposure. [Letter]. *Environ. Health Perspect.*, *107*(5), A236–237.
26. U.S. EPA. (2001). *Water quality criterion for the protection of human health: Methylmercury chapter 4: Risk assessment for methylmercury*. Washington, DC: EPA Office of Sciences and Technology, Office of Water.
27. Ginsberg, G. L., & Toal, B. F. (2000). Development of a single-meal fish consumption advisory for methylmercury. *Risk Analysis*, *20*, 41–47.
28. Swartout, J., & Rice, G. (2000). Uncertainty analysis of the estimated ingestion rates used to derive the methylmercury reference dose. *Drug Clin. Toxicol.*, *23*(1), 293–306.
29. Vahter, M., Akesson, A., Lind, B., Bjors, U., Schutz, A., & Berglund, M. (2000). Longitudinal study of methylmercury and inorganic mercury in blood and urine of pregnant and lactating women, as well as in umbilical cord blood. *Environmental Research*, *84*, 186–194.
30. Sørensen, N., Murata, K., Budtz-Jørgensen, E., Weihe, P., & Grandjean, P. (1999). Prenatal methylmercury exposure as a cardiovascular risk factor at seven years of age. *Epidemiology*, *10*, 370–375.
31. Salonen, J. T., Seppanen, K., Nyyssonen, K., Korpela, H., Kauhanen, J., Kantola, M., Tuomilehto, J., Esterbauer, H., Tatzber, F., & Salonen, R. (1995). Intake of mercury from fish, lipid peroxidation, and the risk of myocardial infarction and coronary, cardiovascular, and any death in Eastern Finnish men. *Circulation*, *91*(3), 645–655.
32. Kinjo, Y., Higashi, H., Nakano, A., Sakamoto, M., & Sakai, R. (1993). Profile of subjective complaints and activities of daily living among current patients with Minamata disease after 3 decades. *Environmental Research*, *63*(2), 241–251.
33. Fukuda, Y., Ushijima, K., Kitano, T., Sakamoto, M., & Futatsuka, M. (1999). An analysis of subjective complaints in a population living in a methylmercury-polluted area. *Environmental Research*, *81*, 100–107.
34. Spyker, J. M. (1975). Assessing the impact of low level chemicals on development: Behavioral and latent effects. *Federation Proceedings*, *34*(9), 1835–1844.
35. Newland, C. M., & Rasmussen, E. B. (2000). Aging unmasks adverse effects of gestational exposure to methylmercury in rats. *Neurotoxicology and Teratology*, *22*, 819–828.
36. Rice, D. C. (1998). Age-related increase in auditory impairment in monkeys exposed in utero plus postnatally to methylmercury. *Toxicological Sciences*, *44*(2), 191–196.
37. Rice, D. C. (1989). Delayed neurotoxicity in monkeys exposed developmentally to methylmercury. *Neurotoxicology*, *10*, 645–650.
38. Rice, D. C., & Hayward, S. (1999). Comparison of visual function at adulthood and during aging in monkeys exposed to lead or methylmercury. *Neurotoxicology*, *20*, 767–784.

STATEMENT OF DR. GARY MYERS, PEDIATRIC NEUROLOGIST AND PROFESSOR, UNIVERSITY OF ROCHESTER

Thank you for the opportunity to present the views of our research group on the health effects of methylmercury (MeHg) exposure. My name is Gary Myers. I am a pediatric neurologist and professor at the University of Rochester in Rochester, New York and one member of a large team that has been studying the human health effects of MeHg for nearly 30 years. For nearly 20 years our group has specifically studied the effects of prenatal MeHg exposure from fish consumption on child development.

MERCURY POISONINGS

In the 1950's, massive industrial pollution for over two decades in Japan resulted in high levels of MeHg in ocean fish. Several thousand cases of human poisoning from consuming the contaminated fish were reported. The precise level of human exposure causing these health problems was never determined, but was thought to be high. During that epidemic pregnant women who themselves had minimal or no clinical symptoms of MeHg poisoning delivered babies with severe brain damage manifested by cerebral palsy, seizures and severe mental retardation. This suggested that MeHg crosses the placenta from the mother to the fetus and that the developing nervous system is especially sensitive to its toxic effects.

In 1971–1972 there was an epidemic of MeHg poisoning in Iraq. Unlike the Japanese poisonings, the source of exposure in Iraq was maternal consumption of seed grain coated with a MeHg fungicide. Our research team studied the children of about 80 women who were pregnant during this outbreak. We measured mercury exposure to the fetus using maternal hair, the biomarker that best corresponds to MeHg brain level, and examined the children. We concluded that there was a possibility that exposure as low as 10 ppm in maternal hair might be associated with adverse effects on the fetus, although there was considerable uncertainty in this estimate. This value is over 10 times the average in the United States, but individuals consuming large quantities of fish can achieve this level.

MERCURY FOUND NATURALLY IN THE ENVIRONMENT

Mercury is a natural element in the earth's crust. In aquatic environments, bacteria can convert inorganic mercury to MeHg. Once MeHg enters the food chain, it is bioaccummulated and bioconcentrated. All fish contain small amounts, and predatory fish or mammals such as whales have larger amounts. Most commercial oceanic fish in the United States has <0.5 ppm MeHg in the muscle, but some freshwater fish have >1 ppm. In comparison, contaminated fish in Japan that caused poisoning had up to 40 ppm.

Everyone who consumes fish is exposed to MeHg, and regular fish consumption can lead to hair mercury levels as high as 10 ppm or more. The average hair mercury level in the United States is < 1 ppm. If MeHg does affect the developing brain at such low levels, mothers who consume large amounts of fish would be exposing their babies to this risk.

The hypothesis of our study in the Seychelles was that prenatal MeHg from fish consumption might affect child development. Since millions of people around the world consume fish as their primary source of protein, we decided to investigate the question directly. We initiated the Seychelles Child Development Study in 1983 and began enrolling subjects in a pilot study in 1987. We selected the Seychelles as a sentinel population for the United States for two reasons. First, they consume large amounts of fish. The average mother in our main cohort ate fish with 12 meals per week or over 10 times that of U.S. women. Second, the fish consumed in Seychelles (average mercury content 0.3 ppm) has approximately the same mercury concentration as commercial fish in the United States.

THE SEYCHELLES CHILD DEVELOPMENT STUDY (SCDS)

The SCDS is a collaborative study carried on by researchers at the University of Rochester Medical Center in Rochester, NY and the Ministries of Health and Education in the Republic of the Seychelles. Funding has come from the National Institute of Environmental Health Sciences, the Food and Drug Administration, and the governments of Seychelles and Sweden. The Republic of the Seychelles is an island Nation in the Indian Ocean off the East Coast of Africa.

Our original hypothesis was that prenatal MeHg exposure at levels achieved by regular maternal consumption of fish would be associated with adverse effects on child development that could be detected by clinical examination, or by the use of

developmental tests that have previously been used to study the effects of environmental exposures on child development.

The Seychelles was chosen partly because there is no mercury pollution and many factors that complicate epidemiological studies of low-level exposure are not present. Health care is free, universal and readily available. Prenatal care is nearly 100 percent, the birth rate is high, and the general health of mothers and children is good. Education is free, universal, and starts at age 3.5 years. There is limited emigration and both the people and the government were cooperative and supportive.

Before starting a carefully controlled main study, we carried out a pilot study. We expected to find only subtle effects on children at these levels of exposure. Consequently, it was important to minimize any possibility of bias, so a number of decisions were made before the study began. First, no one in Seychelles including researchers visiting the island would know the exposure level of any child or mother, unless our results indicated that children were at risk from prenatal mercury exposure. Second, children with a known cause of developmental delay (meningitis, very low birth weight, or brain trauma) would not be studied. Third, the tests administered would include tests previously reported to show associations with MeHg exposure, tests used with other toxic exposures, and other tests that might detect subtle changes. Fourth, all testing would be performed within specific age windows to minimize the effect of age on test interpretation. Fifth, results would be adjusted for multiple confounding factors (covariates), including things like socioeconomic status, maternal intelligence and birth weight, which are known to have independent effects on child development and if not accounted for, could bias the results. Sixth, the data analysis plan would be determined before the data were collected to minimize the possibility that the data would be repeatedly analyzed until the anticipated effect was eventually found.

In 1989–90, we enrolled over 700 mothers and children in the SCDS main study. These children were evaluated on five occasions (6, 19, 29, 66 and 107 months of age) during the past 9 years. When the children were about 4 years old their homes were visited and evaluated. The study focused on prenatal exposure. This was measured in the mothers' hair growing during pregnancy. Postnatal exposure was also periodically measured in the children's hair. The exposure of both mothers and children ranged from 1 to 27 ppm, the range of concern. The testing was extensive with over 57 endpoints being evaluated to date.

Through 107 months (9 years) and over 57 primary endpoints, the study has found only three statistical associations with prenatal MeHg exposure. One of these associations was adverse, one was beneficial and one was indeterminate. These results might be expected to occur by chance and do not support the hypothesis that adverse developmental effects result from prenatal MeHg exposure in the range commonly achieved by consuming large amounts of fish. The test results do show associations with factors known to affect child development such as maternal IQ and home environment so there is evidence that the tests are functioning well.

OUR INTERPRETATION OF THE FINDINGS

We do not believe that there is presently good scientific evidence that moderate fish consumption is harmful to the fetus. However, fish is an important source of protein in many countries and large numbers of mothers around the world rely on fish for proper nutrition. Good maternal nutrition is essential to the baby's health. Additionally, there is increasing evidence that the nutrients in fish are important for brain development and perhaps for cardiac and brain function in older individuals.

The SCDS is ongoing and we will continue to report our results. Presently we are examining a new cohort to determine specific nutrients that might influence the effects of MeHg.

Appendix—Not read before the committee, but included in the handout.

Because of the public health importance of the question being studied by the SCDS, the potential exists for differing opinions of scientific findings to become highly politicized. The SCDS has received only one published criticism (JAMA, 280:737, 1998), but other points have been raised at conferences. These questions are addressed here individually.

• Why did the SCDS measure mercury in the hair rather than in the cord blood? Hair mercury was used because it is the standard measure used in nearly all other studies of this question. Mercury is thought to enter the hair and brain in a similar fashion. Hair was also chosen because hair has been shown to follow blood concentrations longitudinally, and samples of hair can recapitulate the entire period of exposure, in this case the period of gestation. As part of our research we have shown that hair levels reflect levels in the target tissue, brain. Measuring mercury in blood

requires correction for the red blood cell volume (hematocrit) since the mercury is primarily in red blood cells and reflects only very recent exposure. It can also vary if recent meals with high mercury content are consumed.

• Did the SCDS use subjects whose mercury values were too low to detect an association? No, the study's goal was to see if the children of women who consume fish regularly were at risk for adverse developmental effects from MeHg. Women in Seychelles eat fish daily and represent a sentinel population with MeHg levels 10 times higher than U.S. women. Because of higher levels of exposure, their children should be more likely to show adverse effects if they are present. These children show no adverse effects through 9 years of age suggesting that eating ocean fish, when there is no local pollution, is safe. However, we cannot rule out an adverse effect above 12–15 ppm since we had too few cases to substantiate a statistical association if one really existed.

• Did the SCDS use the best tests available to detect developmental problems? Yes, the SCDS used many of the same neurodevelopmental and neuropsychological tests used in other developmental studies. These tests are deemed to be excellent measures for determining development at the ages studied. The tests examined specific domains of children's learning and were increasingly sophisticated as the children become older.

• Did the SCDS find expected associations between development and birth weight, socioeconomic factors, and other covariates? Yes, expected relationships with many covariates such as maternal IQ, family socioeconomic status and the home environment were found, indicating that our tests were sensitive to developmental differences.

• Did the removal of statistical outliers in the analysis bias the study? No. It is standard practice among statisticians to remove statistical outliers. Outliers are values that are inconsistent with the statistical model employed to analyze the data. Every statistical analysis depends on a model, and every statistical model makes assumptions about the statistical (distributional) properties of the data that must be satisfied if the results of the analysis are to be interpreted correctly. Sound statistical practice requires that the necessary assumptions be checked as part of the statistical analysis. Examination of outliers constitutes one of these checks. Statistical outliers are defined by the difference between the actual test score for a child and the value predicted by the statistical model. Small numbers of such outliers occurred in test scores for children with widely varying MeHg exposures. The results of all analysis were examined both before as well as after the removal of outliers. For analyses in the main study the removal of statistical outliers did not change the conclusions.

• What about the Faroe Islands study where prenatal MeHg exposure was reported to adversely affect developmental outcomes? There are substantial differences between the Faroe Islands and Seychelles studies. The exposure in the Faroe Islands is from consuming whale meat and there is also concomitant exposure to PCBs and other neurotoxins. There are also differences in the measurement of exposure and the approach to statistical analysis. The Faroe Islands study reported associations between cord blood mercury levels and several tests. After statistical analysis they attributed the associations to prenatal MeHg exposure. Scientific studies are frequently open to different interpretations and some scientists do not agree with the researchers' interpretation. We believe the Seychelles study of individuals consuming fish more closely approximates the U.S. situation.

• Are the children in Seychelles too developmentally robust to find the effects of MeHg if they are present? No, the children in Seychelles tested similar to U.S. children on nearly all measures apart from motor skills where they were more advanced. There is no reason to think that they are too robust to show the effects of prenatal MeHg exposure if any are present.

• Are children in Seychelles exposed to PCBs or other food-born toxins that might have confounded the results? No, sea mammals are not consumed in Seychelles and measured PCBs in the children's blood were low.

• Should data from the Seychelles be considered interim? Maybe. Among developmental studies, a 9-year followup is considered very long and should be adequate to identify associations with most toxic exposures. However, very subtle effects can be more readily tested in older individuals and there is evidence from experimental animals that some effects of early mercury exposure may not appear until the animal ages.

Original Contributions

Effects of Prenatal and Postnatal Methylmercury Exposure From Fish Consumption on Neurodevelopment

Outcomes at 66 Months of Age in the Seychelles Child Development Study

Philip W. Davidson, PhD; Gary J. Myers, MD; Christopher Cox, PhD; Catherine Axtell, PhD; Conrad Shamlaye, MB, ChB; Jean Sloane-Reeves, MS; Elsa Cernichiari, MS; Larry Needham, PhD; Anna Choi, MS; Yining Wang, PhD; Maths Berlin, MD; Thomas W. Clarkson, PhD

Context.—Human neurodevelopmental consequences of exposure to methylmercury (MeHg) from eating fish remain a question of public health concern.

Objective.—To study the association between MeHg exposure and the developmental outcomes of children in the Republic of Seychelles at 66 months of age.

Design.—A prospective longitudinal cohort study.

Participants.—A total of 711 of 779 cohort mother-child pairs initially enrolled in the Seychelles Child Development Study in 1989.

Setting.—The Republic of Seychelles, an archipelago in the Indian Ocean where 85% of the population consumes ocean fish daily.

Main Outcome Measures.—Prenatal and postnatal MeHg exposure and 6 age-appropriate neurodevelopmental tests: the McCarthy Scales of Children's Abilities, the Preschool Language Scale, the Woodcock-Johnson Applied Problems and Letter and Word Recognition Tests of Achievement, the Bender Gestalt test, and the Child Behavior Checklist.

Results.—The mean maternal hair total mercury level was 6.8 ppm and the mean child hair total mercury level at age 66 months was 6.5 ppm. No adverse outcomes at 66 months were associated with either prenatal or postnatal MeHg exposure.

Conclusion.—In the population studied, consumption of a diet high in ocean fish appears to pose no threat to developmental outcomes through 66 months of age.

JAMA. 1998;280:701-707

From the University of Rochester School of Medicine and Dentistry, Rochester, NY (Drs Davidson, Myers, Cox, Axtell, Wang, and Clarkson, and Mss Reeves, Cernichiari and Choi); the Republic of Seychelles Ministry of Health, Victoria, Mahé, Seychelles (Dr Shamlaye); the Centers for Disease Control and Prevention, Atlanta, Ga (Dr Needham); and The University of Lund, Lund, Sweden (Dr Berlin).

Reprints: Philip W. Davidson, PhD, Strong Center for Developmental Disabilities, Box 671, University of Rochester School of Medicine and Dentistry, 601 Elmwood Ave, Rochester, NY 14642 (e-mail: pdavidson@cc.urmc.rochester.edu)

INORGANIC MERCURY (Hg) discharged into lakes, rivers, and oceans is converted to methylmercury (MeHg) by microorganisms and bioaccumulated up the aquatic food chain.[1] Concern about the potential public health threat from MeHg arose in the United States in the early 1970s when elevated concentrations were found in fish in the Great Lakes. Today, recreational fishing is restricted in many states and Food and Drug Administration guidelines regulate interstate commerce of fish because of their MeHg content.

For editorial comment see p 737.

Mass health disasters in Minamata and Niigata, Japan, caused by consumption of fish highly contaminated with MeHg from an industrial source,[2,3] and in Iraq following consumption of bread containing MeHg fungicide,[4-6] confirmed that MeHg was neurotoxic and that the prenatal period was the most sensitive stage of the life cycle. For example, severe exposures in Iraq (up to 674 ppm of mercury in hair) were associated with microcephaly, seizures, mental retardation, and cerebral palsy. The Iraq outbreak also resulted in less severe outcomes typified by developmental delays and abnormal results of neurological examinations. A dose-response analysis suggested that effects may occur at maternal hair concentrations of mercury as low as 10 ppm,[5,6] although there was considerable uncertainty in this estimate. This compares with an average in the US population of 1 ppm or less.[7]

All fish contain MeHg. Frequent consumption of ocean fish can lead to MeHg levels in excess of 10 ppm and as high as 50 ppm in hair.[1] Epidemiological studies[8-10] on populations consuming fish where Hg was biologically methylated failed to find

clinical cases of MeHg poisoning. The possibility that prenatal MeHg exposure from maternal consumption of a fish diet may be associated with subtle changes in children's cognitive and neurological development has been examined in these studies with inconclusive results.[1,11]

We have followed longitudinally[12,13] a large inception cohort of mother-child pairs in the Republic of Seychelles, a westernized archipelago in the middle of the Indian Ocean where 85% of the population consumes marine fish daily.[14] This article presents the results of the neurodevelopmental examination of the Seychelles Child Development Study (SCDS) cohort at 66 months, an age at which neuropsychological tests may be given that are sensitive enough to assess potential associations between developmental outcomes and MeHg dietary exposure. Our results also include examination of the role of postnatal exposure from fish consumption.

METHODS

Subjects

The cohort consisted of 711 mother-child pairs living in the Republic of Seychelles, representing 91% of the 779 pairs originally enrolled in the SCDS main study.[11] Informed consent was obtained from the caregiver of every participating child. The protocol was approved by human subjects review boards at the University of Rochester, Rochester, NY, and the Ministry of Health, Victoria, Mahé, Republic of Seychelles, before enrollment began. The sample size was sufficient to detect a 5.7-point difference on any test with a mean (SD) of 100 (16) between low (0-3 ppm) and high (>12 ppm) MeHg groups for a 2-sided test ($\alpha = .05$ at 80% power).

Twenty-eight mother-child pairs were excluded because of medical problems that might seriously affect development.[13] An additional 16 pairs had insufficient maternal hair available to accurately recapitulate prenatal exposure, and 24 did not return for testing at 66 months.

Demographic characteristics of the Seychelles and the cohort were reported earlier.[15] At enrollment, the mothers reported eating an average of approximately 12 marine fish meals per week. Sea mammals are not consumed in the Seychelles. We have previously documented that lead levels in whole blood are less than 0.48 µmol/L (10 µg/dL) in a representative group of Seychellois children and mothers.[16]

Procedure Test Battery

Each child was evaluated at 66 months (±6 months) at a child development center staffed by a team of specially trained Seychellois nurses blinded to MeHg levels and the results of testing during previous visits. Five children were tested between 72 months and 79 months of age. All evaluations were conducted between July 1994 and October 1995. The test battery assessed multiple developmental domains[17] and was similar to those used to demonstrate adverse developmental effects of exposure to lead[18] and polychlorinated biphenyls (PCBs)[19] and to those used in earlier studies to measure MeHg exposure effects.[9,20] The tests are sufficiently sensitive and accurate to detect neurotoxicity in the presence of a number of confounding factors.[21]

The test battery included the following 6 primary measures: (1) the General Cognitive Index (GCI) of the McCarthy Scales of Children's Abilities[22] to estimate cognitive ability; (2) the Preschool Language Scale[23] (PLS) total score to measure both expressive and receptive language ability; (3) the Letter and Word Recognition and (4) the Applied Problems subtests of the Woodcock-Johnson (W-J) Tests of Achievement[24] to measure reading and arithmetic achievement; (5) the Bender Gestalt test[25] to measure visual-spatial ability; and (6) the total T score from the Child Behavior Checklist (CBCL)[26] to measure the child's social and adaptive behavior. The CBCL questionnaire was completed by each child's primary caregiver. All tests were given in Creole, the language spoken by 98% of Seychellois at home.

Pure tone hearing thresholds were tested using a portable audiometer. Caregiver IQ was determined using the Raven Standard Progressive Matrices, a nonverbal test designed to minimize the effects of culture on measurement of IQ.[27] When the children were between 42 and 56 months of age, the Home Observation for Measurement of the Environment (HOME) Inventory for Families of Preschool Age Children[28] was administered during home visits. Following procedures described earlier,[17] on-site test administration reliability was assessed by an independent scorer; percentage of disagreement ranged from 0% to 8%. Mean intraclass correlations for interscorer reliability were 0.96 to 0.97. Reliability of final scoring in Rochester was conducted by rescoring a sample of tests; the mean intraclass correlation coefficient was 0.96.

Mercury Exposure

Prenatal exposure was assessed by measuring the concentration of total mercury (THg) in a segment of maternal hair representing growth during pregnancy. Total Hg in maternal hair during pregnancy correlates well with blood levels of MeHg[1] and with THg levels in fetal brain.[29] Methylmercury accounts for over 80% of the THg in hair samples collected from fish-eating populations.[16,29] Maternal hair levels of THg have been the biological indicator of choice in nearly all previous epidemiological studies of fetal exposure to MeHg. There is considerable variation in the relationship between hair and blood THg in different individuals. However, the key relationship, that of hair levels and brain levels, may not show the same variability.[29] Postnatal exposure was determined by measuring THg at 66 months of age from a 1-cm segment of the child's hair nearest the scalp. This age was chosen because it was coincident with the age of testing and all children were postweaning and eating a fish diet. In 25 children, hair samples taken during the HOME administration (at 48 months of age) were used to determine postnatal MeHg exposure because their hair sample at 66 months of age was insufficient for analysis.

Mercury in Fish

Total Hg was analyzed in 5 or more samples of species of fish caught and consumed in the Seychelles, including yellow fin tuna (*Thunnus albacares*), Indian mackerel (*Rastrelleger kanagurta*), brown spotted grouper (*Epinephelus chlorustigma*), green jobfish (*Aprion virescens*), bonito (*Euthynnus affinis*), bludger (*Carangoides gymnostethus*), and spangled emperor (*Lethrinus nebulosus*).

Maternal and Child Mercury Analysis

The analysis of hair samples and fish homogenates for THg and inorganic Hg was done by cold vapor atomic absorption spectrometry. The analysis technique and quality control procedure are given elsewhere.[16]

PCBs in Blood

Polychlorinated biphenyls are known to be present in some ocean fish and may be associated with developmental delays in children. Levels of PCBs in serum of 49 of the children at 66 months of age were analyzed at the laboratories of the US Centers for Disease Control and Prevention, Atlanta, Ga. The analytic method for measuring the PCBs involved deproteinization of the serum with formic acid, elution through a column containing octadodecyl (C18) packing material, elution through a column containing Florisil packing material, concentration of the organic eluents, and analysis by dual capillary column–gas chromatography with electron capture detection.[30]

Table 1.—Neurodevelopmental Test Means (SDs) by Prenatal Exposure Levels*

Test	Subgroupings by Prenatal Total Mercury Exposure Level, Mean (SD), ppm				
	≤3 (Mean, 2-0) (n = 159)	>3-6 (Mean, 4.5) (n = 206)	>6-9 (Mean, 7.4) (n = 156)	>9-12 (Mean, 10.3) (n = 95)	12-26.7 (Mean, 15.3) (n = 95)
McCarthy Scales of Children's Abilities General Cognitive Index	94.0 (12.3)	93.8 (13.1)	94.3 (13.8)	92.4 (11.6)	95.9 (12.6)
Preschool Language Scale total score†	69.6 (6.7)	69.6 (6.7)	69.6 (6.8)	70.2 (6.3)	72.1 (6.6)
Bender Gestalt errors‡	10.2 (3.9)	10.4 (3.7)	10.0 (4.1)	10.5 (3.7)	9.4 (3.9)
Woodcock-Johnson Tests of Achievement					
Letter and Word Recognition	76.1 (10.8)	77.6 (11.1)	76.9 (9.9)	76.1 (9.9)	77.7 (10.9)
Applied Problems	85.6 (17.2)	87.3 (17.7)	87.3 (17.5)	87.0 (18.0)	90.1 (17.9)
Child Behavior Checklist total T score§	60.4 (9.7)	59.7 (10.5)	59.3 (10.2)	59.3 (9.8)	59.7 (8.8)

*N = 711.
†The Preschool Language Scale raw scores were used since the test and its norms were based on English.
‡The Koppitz scoring method was used. The score represents the number of indicator errors.
§The Child Behavior Checklist yields a percentile (T) score. The threshold for abnormality is beyond the 75th percentile.

Table 2.—Neurodevelopmental Test Means (SDs) by Postnatal Exposure Level*

Test	Subgroupings by Postnatal Total Mercury Exposure Level, Mean (SD), ppm				
	≤3 (Mean, 2.2) (n = 73)	>3-6 (Mean, 4.6) (n = 299)	>6-9 (Mean, 7.4) (n = 213)	>9-12 (Mean, 10.2) (n = 76)	12-25.8 (Mean, 14.9) (n = 47)
McCarthy Scales of Children's Abilities General Cognitive Index	90.6 (14.3)	94.0 (12.8)	94.7 (12.5)	95.8 (13.5)	93.0 (10.5)
Preschool Language Scale total score†	68.8 (6.8)	69.6 (6.7)	70.7 (6.5)	70.6 (7.8)	70.1 (4.8)
Bender Gestalt errors‡	10.4 (4.0)	10.2 (3.9)	10.1 (3.6)	10.0 (4.4)	10.1 (4.0)
Woodcock-Johnson Tests of Achievement					
Letter and Word Recognition	74.7 (11.0)	77.8 (11.3)	76.7 (9.5)	76.8 (10.9)	75.6 (9.0)
Applied Problems	85.6 (18.1)	86.7 (17.6)	88.0 (18.2)	90.0 (16.0)	85.0 (16.3)
Child Behavior Checklist total T score§	57.6 (9.4)	60.1 (10.0)	60.2 (10.7)	59.3 (9.3)	59.5 (7.6)

*N = 708.
†The Preschool Language Scale raw scores were used since the test and its norms were based on English.
‡The Koppitz scoring method was used. The score represents the number of indicator errors.
§The Child Behavior Checklist yields a percentile (T) score. The threshold for abnormality is beyond the 75th percentile.

Statistical Analysis

The effect of prenatal and postnatal MeHg on each outcome variable was adjusted for covariates, specified as part of the study design, and selected because of their potential to bias the assessment of the association between Hg and outcome.[17] Covariates associated with the child included birth weight, birth order, sex, history of breast-feeding, hearing status, and the child's medical history. Covariates associated with the mother and family included maternal age, maternal smoking during pregnancy, maternal alcohol consumption during pregnancy, maternal medical history, caregiver intelligence, language spoken in the home, Hollingshead socioeconomic status (SES), and HOME score. Two multiple linear regression analyses (with 2-tailed significance tests using a significance level of $P \leq .05$) including both prenatal and postnatal THg were performed for each of the 6 primary measures. The first involved all covariates (full model) and the second included only covariates believed to most likely influence child development in the Seychelles (reduced model), including sex, birth weight, child's medical history, maternal age, HOME score, caregiver IQ, SES, and hearing status. Each full and reduced model was run both with and without THg by sex interaction terms to test the hypothesis that males and females have different THg slopes.

All models were examined for statistical outliers and influential points.[31] Each model was run first with outliers, then without outliers, and the results were compared. All of the results were essentially the same with or without outliers. The regression analyses for all 6 primary measures were also repeated without influential points to determine whether the original results were dependent on such points.[31] The final analysis included influential points that were not also outliers. The results without influential points were consistent with the original analysis.

Secondary analyses tested the hypothesis that associations between developmental outcomes and THg exposure might be nonlinear. All regression analyses were repeated, first using the log of the prenatal and postnatal THg values, then classifying THg variables into 5 groups each for prenatal and postnatal exposure.

RESULTS

Mercury in Fish

A total of 350 samples of fish were analyzed. The median THg for each of the 25 species ranged from 0.004 ppm to 0.75 ppm, with most medians in the range of 0.05 to 0.25 ppm. These levels are comparable with fish in the US market. The lowest levels occurred in reef fish. Methylmercury accounted for over 90% of the THg in 34 fish homogenates analyzed by gas chromatography–atomic fluorescent detection. This finding confirms many previous observations.[32]

PCBs in Blood

Twenty-eight PCB congeners ranging from congener 28 to 206 were measured in each serum sample. All samples had no detectable levels of any PCB congeners. The detection limit for the PCB analysis was 0.2 ng/mL. These results are typical for persons with no known exposure to PCBs.[33]

Mercury Exposure

The mean (SD) maternal hair level of THg during pregnancy was 6.8 (4.5) ppm (n = 711), and the mean child hair level at 66 months was 6.5 (3.3) ppm (n = 708). The ranges (maternal hair, 0.5-26.7 ppm; child hair, 0.9-25.8 ppm) were sufficient to test for exposure effects using regression analysis. Maternal and child THg concentrations were not highly associated Pearson $r = 0.15$, n = 708, $P < .001$) as observed by others.[20] The exposure levels found in the Seychelles are typical of populations that depend on fish as a major dietary source of protein and calories.[34]

Test Performance

Tables 1 and 2 show test score means and SDs for prenatal and postnatal exposure levels. These results are similar to what would be expected from a healthy, well-developing US population. No test indicated a deleterious effect of MeHg exposure. Four of the 6 measures showed better scores in the highest MeHg groups compared with lower groups for both prenatal and postnatal exposure.

Regression Analyses

Primary Analyses.—The models with all covariates (full models) and limited covariates (reduced models) were both significant (ie, each model was able to describe the data) and yielded similar results for every measure. The THg by sex interaction test for prenatal exposure was not significant in any regression model. The THg by sex interaction was significant for postnatal exposure for the Bender Gestalt test; hence, we report the results of the reduced model with both interaction terms included. For all

Table 3.—Multiple Regression Analysis of 66-Month Test Scores: Parameter Estimates for Reduced Models With Outliers Removed*

	Regression Coefficients (SEs)†					
				Woodcock-Johnson Tests of Achievement‡		
Parameter Estimate	McCarthy GCI‡	PLS Total Score‡	Bender Gestalt Errors§	Applied Problems	Letter and Word Recognition	CBCL Total T Score‡
Maternal MeHg	−0.057 (0.10)	0.13 (0.057)	0.04 (0.05)	0.11 (0.14)	0.02 (0.09)	−0.11 (0.09)
Child MeHg	0.26 (0.14)	0.18 (0.075)	−0.16 (0.06)	0.36 (0.18)	0.15 (0.12)	−0.02 (0.11)
Sex, female	2.5 (0.92)‖	0.97 (0.51)	−0.14 (0.73)	3.99 (1.24)	1.68 (0.78)	0.002 (0.76)
Sex × maternal MeHg	...	...	...	...	...	...
Female subjects × maternal MeHg	...	...	−0.70 (0.046)	...	...	...
Male subjects × maternal MeHg	...	...	0.042 (0.047)	...	...	...
Sex × child MeHg	...	...	...	...	...	...
Female subjects × child MeHg	...	...	0.09 (0.064)	...	...	...
Male subjects × child MeHg	...	...	−0.16 (0.06)	...	...	...
Birth weight	1.4 (0.98)	0.84 (0.54)	−0.34 (0.31)	2.98 (1.31)‖	1.07 (0.82)	−0.08 (0.81)
Child's medical history¶	−0 65 (2.69)	−0.72 (1.40)	−1.09 (0.83)	2.87 (3.51)	3.57 (2.25)	−2.06 (2.16)
Maternal age	0 054 (0.09)	0.008 (0.47)	−0.02 (0.03)	−0.003 (0.11)	0.01 (0.07)	−0.25 (0.07)
HOME (*P* for overall test)‖	< 01	<.01	<.01	>.01 and ≤.05	<.01	<.01
Low (0-31)	−4.4 (1.22)	−3.16 (0.67)	1.39 (0.38)	−4.69 (1.63)	−4.53 (1.02)	2.79 (1.0)
Medium (>31-35)	−3.2 (1.19)	−1.85 (0.65)	00.75 (0.38)	−3.52 (1.59)	−4.45 (1.0)	2.48 (0.98)
SES (*P* for overall test)‖	.11	.15	.48	>.01 and ≤.05	≤.01	.14
Unskilled	−4.1 (1.68)	−2.01 (0.93)	0.72 (0.53)	−6.44 (2.26)	−4.73 (1.4)	3.0 (1.37)
Skilled	−2.9 (1.57)	−1.14 (0.86)	0.41 (0.5)	−4.57 (2.12)	−4.31 (1.31)	2.74 (1.29)
Semiprofessional	−2.41 (1.65)	−0.76 (0.91)	0.14 (0.52)	−3.92 (2.23)	−0.65 (1.38)	1.78 (1.35)
Caregiver IQ (*P* for overall test)‖	>.01 and ≤.05	<.01	<.01	<.01	.78	.59
Lower third	−3.3 (1.26)	−2.22 (0.7)	1.33 (0.4)	−5.55 (1.69)	−0.65 (1.06)	1.06 (1.03)
Middle third	−1.61 (1.14)	−1.06 (0.62)	0.79 (0.36)	−2.23 (1.52)	−0.59 (0.96)	0.57 (0.94)
Hearing level (*P* for overall test)‖	.96	.87	.09	.03	.52	.99
0-25 dB	−0.94 (3.53)	−0.02 (2.22)	2.45 (1.22)	−6.34 (4.92)	−0.44 (3.10)	0.20 (2.9)
26-35 dB	−1.2 (4.03)	−0.02 (2.48)	1.60 (1.36)	−3.81 (5.56)	−2.38 (3.50)	0.43 (3.3)

*GCI indicates General Cognitive Index; PLS, Preschool Language Scale; CBCL, Child Behavior Checklist; MeHg, methylmercury; HOME, Home Observation for Measurement of the Environment; and SES, socioeconomic status. Ellipses indicate data not applicable.
†Slope-test point/ppm of methylmercury.
‡Model without interaction.
§Model with test for methylmercury by sex interaction.
‖Slopes are provided for the difference between each level of the covariate and the reference (highest) level. Indications of statistical significance are provided for these individual effects as well as for overall effects.
¶Reference group is positive history.

other analyses, we report the results for reduced models without THg by sex interactions.

The regression coefficients for all variables in the 6 sets of analyses are shown in Table 3. Figure 1 shows partial residual plots (end points adjusted for covariates) for prenatal and postnatal exposure for the McCarthy GCI, the PLS total score, and the W-J Applied Problems test score. For the McCarthy GCI analysis, the model (F [15, 628] = 4.41, $P<.001$, $R^2 = 0.10$) indicated that slopes for both THg exposures did not differ from 0 (P = .59 and .06 for prenatal and postnatal exposure, respectively). For the PLS analysis, the model (F [15, 590] = 6.25, $P<.001$, $R^2 = 0.14$) showed effects of both prenatal and postnatal THg exposure (P = .02 for both), but the effects were very small and in a direction of enhanced performance. The total increase in scores across the entire range of THg exposures was less than 4.5 points. The W-J Applied Problems model (F [15, 625] = 5.38, $P<.001$, R^2 = 0.11) indicated that the slope for prenatal exposure did not differ significantly from 0 (P = .41). There was a significant beneficial postnatal exposure effect (P = .05), but no evidence for an adverse effect. The postnatal exposure slope for the W-J Applied Problems test was 0.36 ppm, representing a 9.7-point increase over the full exposure range, or a 10% improvement in performance.

Figure 2 shows partial residual plots for the Bender Gestalt test, separating male from female subjects. The model (F [17, 613] = 4.53, $P<.001$, $R^2 = 0.11$] showed no significant association with prenatal exposure. The interaction of postnatal THg with sex was significant (P = .004). The slightly positive slope for female errors was not significant (P = .14). The regression coefficient of −0.16 ppm for male subjects was significant (P = .009), resulting in a reduction of 4.3 errors over the entire exposure range (a 40% performance improvement given that the average score for the lowest-exposure group was about 10 errors).

Although the models for the W-J Letter and Word Recognition achievement score and the CBCL yielded significant overall F statistics, none of these was significantly associated with either prenatal or postnatal exposure.

The data for covariate effects shown in Table 3 indicate that test scores were frequently influenced by sex (female subjects scored higher than male subjects), and were directly related to SES, quality of home stimulation, and caregiver IQ, as would be expected in westernized cultures. These data also indicate that performance by Seychellois children on these tests was similar to what would be expected of US children.

Table 4 gives the partial R^2 values for the effects of prenatal and postnatal THg exposure on each developmental measure and the 95% confidence intervals for the effect of a 10-ppm increase in hair THg concentration. The small partial R^2 values shown in Table 4 indicate that THg exposure accounted for little of the variance associated with each test. For the McCarthy GCI and the CBCL, the magnitude of the negative lower confidence limit might be of clinical significance, at least over a greater range of hair THg levels. However, the inclusion of 0 in the confidence intervals shows

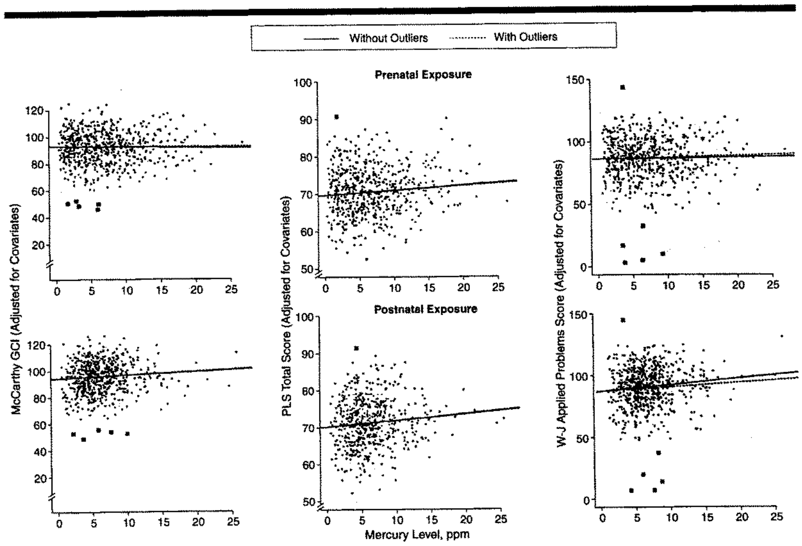


Figure 1.—Partial residuals for prenatal and postnatal exposure. The measures are the McCarthy Scales of Children's Abilities General Cognitive Index (GCI), the Preschool Language Scale (PLS) total score, and the Woodcock-Johnson (W-J) Applied Problems subtest. Each test score was adjusted for all reduced model predictors except the exposure value used in the plot. For graphical representation, the overall mean test score was added to the resulting partial residual. The slope of the line in the plot is the regression coefficient for the multiple regression model. Slopes are shown for the model with and without outliers. Black squares indicate outliers.

that the effects are not statistically significant.

Secondary Analysis.—Log transformations of the 2 THg variables did not alter the direction of any effects. In the categorical analyses, test scores for children with an MeHg exposure level greater than 12 ppm were not significantly different than for children with exposure levels of 3 ppm or less.

COMMENT

Results from this study are relevant for the United States and other countries with similar dietary intake of fish. The major source of MeHg in the Seychelles is ocean fish and the average fish levels are similar to those on the US market. Seychellois MeHg levels are 10 to 20 times higher than in the United States because the Seychellois consume more fish, not because they eat a few fish with abnormally high MeHg levels. Thus, any potential adverse effects of MeHg in fish should be detected in the Seychelles long before such effects would be seen in the United States. Our findings confirm our earlier report in which no adverse developmental effects were found in toddlers following prenatal MeHg exposure.[12] Our data extend preliminary results from the SCDS Pilot Study,[35] involving a less statistically powerful, less well-controlled developmental evaluation of 217 Seychellois children at 66 months using most of the same measures but lacking many of the covariates used here.

We applied multiple regression analysis to our data as has been done in other studies on the effects of mercury,[9-13] lead,[18,36] or PCBs[19] on child development. Although our models showed no negative associations between MeHg and outcome scores, our test procedures did detect other factors known to be associated with child development. The HOME test indicated that the quality of the home environment had a substantial impact on child performance, significantly affecting the results of all tests. The SES of the family, the caregiver's IQ, and the child's sex were all found to have an influence on performance scores, in keeping with the literature on child development.[37,38] These results increase confidence in the sensitivity of our tests to child development functions. Postnatal exposure to MeHg at 66 months of age was associated with a small but statistically significant increase on several developmental outcomes. Our hypothesis did not predict positive effects, since there are no reasons to suppose that such effects are associated with exposure to MeHg. However, MeHg levels in hair are known to correlate closely with fish intake, and other factors or agents associated with fish, such as omega-3 fatty acids, may have beneficial effects. A large cohort study under way in the Faroe Islands found enhancement of developmental milestones in suckling infants exposed to MeHg in breast milk.[39] They suggested that MeHg levels in the infant were a surrogate for the length of breast-feeding, which is reported to have a positive association with developmental outcomes.[40]

In contrast with the conclusions from this and our earlier studies of the main Seychellois cohort,[12,13] the Faroe Islands study found evidence of cognitive deficits associated with prenatal exposure to THg when children were tested at 7

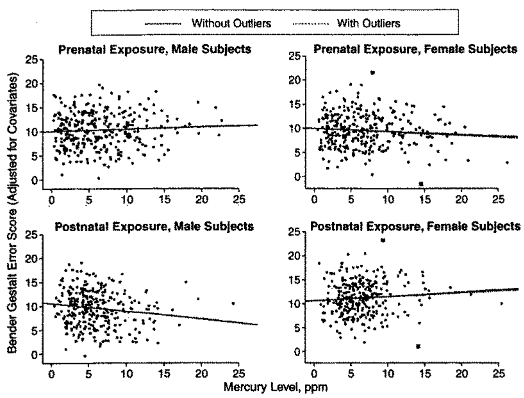


Figure 2.—Partial residuals for prenatal and postnatal exposure. The measures are Bender Gestalt error scores for male and female subjects. Each test score was adjusted for all reduced model predictors except the exposure value used in the plot. For graphical representation, the overall mean test score was added to the resulting partial residual. The slope of the line in the plot is the regression coefficient for the multiple regression model. Slopes are shown for the model with and without outliers. Black squares indicate outliers.

Table 4.—Partial R^2 Values for Total Mercury and 95% CIs for a 10-ppm Increase in Both Prenatal and Postnatal Total Mercury Hair Levels From the Regression Analysis of 66-Month End Points*

	Partial R^2 Value, %		95% CIs for a 10-ppm Increase	
Developmental End Point	Prenatal	Postnatal	Prenatal	Postnatal
McCarthy GCI	0.047	0.55	(−2.58 to 1.45)	(−0.12 to 5.29)
PLS total score	0.91	0.96	(0.22-2.41)	(0.34-3.24)
Woodcock-Johnson Tests of Achievement				
Letter and Word Recognition	0.0062	0.26	(−1.51-1.86)	(−0.79 to 3.75)
Applied Problems	0.11	0.61	(−1.57 to 3.84)	(0.02-7.23)
CBCL total T score	0.24	0.0058	(−2.72 to 0.61)	(−2.43 to 2.00)
Bender Gestalt error score				
Male subjects	0.13	1.11	(−0.50 to 1.34)	(−1.58 to 0.18)
Female subjects	0.38	0.35	(−2.76 to −0.41)	(−0.29 to 2.18)

*Data are from reduced models without interaction (except for those for the Bender Gestalt error score) and with outliers removed. CI indicates confidence interval; GCI, General Cognitive Index; PLS, Preschool Language Scale; and CBCL, Child Behavior Checklist.

years of age.[20] Important differences between the 2 populations may explain the divergent outcomes (eg, nutritional practices, housing, and lifestyle). However, the major difference between our study and the Faroe Islands study is the source of exposure. Ocean fish are the source of MeHg in the Seychelles, whereas pilot whales are the predominant source in the Faroe Islands.[20] The average MeHg level in the meat of pilot whales sampled in the Faroe Islands was 1.6 ppm,[41] approximately 10 times higher than the average level in fish consumed in the Seychelles. Approximately the same level of inorganic Hg is also present in whale meat.[41] In addition, whale blubber is also consumed by the Faroese[41] and is heavily contaminated with fat-soluble pollutants.[42] The average PCB concentration in pilot whale blubber from Faroese waters is elevated (about 30 ppm).[43] In general, fatty tissues of marine mammals in the North Atlantic also contain elevated levels of persistent organochlorine compounds including dibenzofurans and dioxins, DDT and its metabolites, and other pesticides. It is difficult to determine the relative toxicological impact of individual compounds. Some of these contaminants are believed to affect child development.[44] The Faroese study may be relevant to populations consuming large, perhaps episodic, amounts of marine mammals, but its relevance to people consuming ocean fish remains to be established.

A Swedish expert group conducted the first extensive evaluation of human health risks from MeHg in fish in 1971.[45] They concluded the lowest toxic level in hair was 50 ppm in adults. The World Health Organization (WHO) expert group[46] subsequently reaffirmed the Swedish conclusion and applied a safety factor of 10 to cover risks to the most sensitive subgroup of the population, assumed to be those who are prenatally exposed. Thus, 5 ppm in hair was adopted as the international standard for the upper tolerable level of Hg in hair. Subsequent epidemiological studies of human populations prenatally exposed to MeHg from fish have given strong support to the WHO guideline.[8,10,12,13] Our results add further support to the validity of this longstanding guideline.

In summary, the results of extensive performance tests conducted with cohort children at 66 months of age strongly support our findings reported at younger ages. The development of these children is proceeding well without any detectable adverse influence of MeHg. Our results support Egeland and Middaugh's observation[47] that it would be inadvisable to forgo the health benefits of fish consumption to protect against a small risk of adverse effect at the levels of MeHg found in ocean fish on the US market.

This research is supported by grants ES-05497, ES-01247, and ES-07271 from the National Institutes of Health, Bethesda, Md, by a grant from the Food and Drug Administration, US Department of Health and Human Services, Washington, DC, and by the Ministry of Health, Victoria, Mahé, Republic of Seychelles.

References

1. *World Health Organization Environmental Health Criteria 101: Methylmercury.* Geneva, Switzerland: World Health Organization; 1990.
2. Matsumoto H, Koya G, Takeuchi T. Fetal Minamata disease. *J Neuropathol Exp Neurol.* 1965;24:563-574.
3. Takeuchi T. Pathology of Minamata disease. In: Study Group of Minamata Disease, ed. *Minamata Disease.* Kumamoto, Japan: Kumamoto University; 1968:141-228.
4. Marsh DO, Myers GJ, Clarkson TW, Amin-Zaki L, Tikriti S, Majeed M. Fetal methylmercury poisoning: clinical and toxicological data on 29 cases. *Ann Neurol.* 1980;7:348-353.
5. Cox C, Clarkson TW, Marsh DO, Amin-Zaki L, Tikriti S, Myers GJ. Dose-response analysis of infants prenatally exposed to methylmercury: an application of a single compartment model to single-strand hair analysis. *Environ Res.* 1989;31:640-649.
6. Cox C, Marsh DO, Myers GJ, Clarkson TW. Analysis of data on delayed development from the 1971-72 outbreak of methylmercury poisoning in Iraq: assessment of influential points. *Neurotoxicology.* 1995;16:727-730.
7. Smith JC, Allen PV, Von Burg R. Hair methylmercury levels in U.S. women. *Arch Environ Health.* 1997;52:476-480.
8. McKeown-Eyssen G, Reudy J, Neims A. Methylmercury exposures in northern Quebec, II: neurologic findings in children. *Am J Epidemiol.* 1983;118:470-479.
9. Kjellstrom T, Kennedy P, Wallis S, Mantell C.

Physical and Mental Development of Children With Prenatal Exposure to Mercury From Fish, Stage I: Preliminary Tests at Age 4. Solna, Sweden: National Swedish Environmental Protection Board; 1986. Report 3080.

10. Kjellstrom T, Kennedy P, Wallis S, et al. *Physical and Mental Development of Children With Prenatal Exposure to Mercury From Fish, Stage II: Interviews and Psychological Tests at Age 6*. Solna, Sweden: National Swedish Environmental Protection Board; 1989. Report 3642.

11. Marsh D, Clarkson TW, Myers GJ, et al. The Seychelles study of methylmercury exposure and child development: introduction. *Neurotoxicology*. 1995;16:583-596.

12. Myers GJ, Marsh DO, Davidson PW, et al. Main neurodevelopmental study of Seychellois children following in utero exposure to methylmercury from a maternal fish diet: outcome at six months. *Neurotoxicology*. 1995;16:653-664.

13. Davidson PW, Myers GJ, Cox C, et al. Longitudinal neurodevelopmental study of Seychellois children following in utero exposure to methylmercury from maternal fish ingestion outcomes at 19 and 29 months. *Neurotoxicology*. 1995,16·677-688.

14. Bovet P, Perret F, Shamlaye C, et al. The Seychelles Heart Study, II: methods and basic findings. *Seychelles Med Dent J*. 1997;1·8-24

15. Shamlaye CF, Marsh DO, Myers GJ, et al. The Seychelles child development study on neurodevelopmental outcomes following in utero exposure to methylmercury from a maternal fish diet background and demographics. *Neurotoxicology* 1995;16:597-612.

16. Cernichiari E, Toribara TY, Liang L, et al. The biological monitoring of mercury in the Seychelles study. *Neurotoxicology*. 1995;16:613-628.

17. Davidson PW, Myers GJ, Cox C, et al. Neurodevelopmental test selection, administration, and performance in the main Seychelles Child Development Study. *Neurotoxicology* 1995;16:665-676.

18. Needleman HL, Gatsonis CA Low level lead exposure and the IQ of children. a meta-analysis of modern studies. *JAMA*. 1990;275 363-369

19. Jacobson JL, Jacobson SW Intellectual impairment in children exposed to polychlorinated biphenyls in utero. *N Engl J Med*. 1996;335.783-789.

20. Grandjean P, Weihe P, White RF, et al Cognitive deficit in 7-year-old children with prenatal exposure to methylmercury. *Neurotoxicol Teratol*. 1997;19:417-428.

21. Amler RW, Gibertini M, eds. *Pediatric Environmental Neurobehavioral Test Battery*. Atlanta, Ga: Agency for Toxic Substances and Disease Registry, US Dept of Health and Human Services; 1996.

22. McCarthy D. *McCarthy Scales of Children's Abilities*. New York, NY: The Psychological Corp; 1972.

23. Zimmerman I, Steiner V, Pond R. *Preschool Language Scale*. Rev ed. Columbus, Ohio: CE Merrill; 1979.

24. Woodcock R, Johnson M. *Woodcock-Johnson Tests of Achievement*. Allen, Tex: DLM, 1989

25. Koppitz EM. *The Bender Gestalt Test for Young Children*. London, England: Grune & Stratton; 1963.

26. Achenbach TM. *Manual for the Child Behavior Checklist and 1991 Child Behavior Profile*. Burlington: University of Vermont Dept of Psychiatry; 1991.

27. Raven J. *Standard Progressive Matrices*. Cambridge, England: HK Lewis; 1958.

28. Caldwell B, Bradley R. *Home Observation of Measurement of the Environment*. Little Rock: University of Arkansas at Little Rock; 1984.

29. Cernichiari E, Brewer R, Myers G, et al. Monitoring methylmercury during pregnancy: maternal hair predicts fetal brain exposure. *Neurotoxicology*. 1995;16:705-710.

30. Brock JW, Bruse VW, Ashley DL, et al. An improved analysis for chlorinated pesticides and polychlorinate biphenyls (PCBs) in human and bovine sera utilizing solid phase extraction. *J Anal Toxicol*. 1996,20 1-9.

31. Cook RD, Weisberg S. *Residuals and Influence in Regression (Monographs on Statistics and Applied Probability)*. New York, NY: Chapman & Hall; 1982.

32. Bloom NS. On the chemical form of mercury in the edible fish and marine invertebrate tissue. *Can J Fish Aquatic Sci*. 1992;49:1010-1017.

33. Agency for Toxic Substances and Disease Registry. *Toxicological Profile for Polychlorinated Biphenyls (Update)*. Atlanta, Ga: US Dept of Health and Human Services; 1997.

34. Airey D. Total mercury concentrations in human hair from 13 countries in relation to fish consumption and location. *Sci Total Environ* 1983;31. 157-180.

35. Myers DJ, Davidson PW, Cox C, et al. Neurodevelopmental outcomes of Seychellois children sixty-six months after in utero exposure to methylmercury from a maternal fish diet: pilot study. *Neurotoxicology*. 1995;16:639-652.

36. Bellinger D, Leviton A, Waternaux C, Needleman H, Rabinowitz M. Longitudinal analyses of prenatal and postnatal lead exposure and early cognitive development. *N Engl J Med*. 1987;316:1037-1043.

37. Bendersky M, Lewis M. Environmental risk, biological risk, and developmental outcome. *Dev Psychol*. 1994;30:484-494.

38. Bendersky M, Lewis M. Effects of intraventricular hemorrhage and other medical and environmental risks on multiple outcomes at age three years. *Dev Behav Pediatr*. 1995;16:89-96.

39. Grandjean P, Weihe P, White RF. Milestone development in infants exposed to methylmercury from human milk. *Neurotoxicology*. 1995;16:32-34.

40. Lucas A, Morley R, Cole TJ, Lister G, Leeson-Payne C. Breast milk and subsequent intelligence quotient in children born preterm. *Lancet*. 1992;339:261-264.

41. Grandjean P, Weihe P, Jørgensen P, Clarkson T, Cernichiari E, Vider T. Impact of maternal seafood diet on fetal exposure to mercury, selenium, and lead. *Arch Environ Health*. 1992;47:185-195.

42. Sanderson K. Marine mammals and the marine environment. *Sci Total Environ*. 1996;186:1-179.

43. Weihe P, Grandjean P, Debes F, White R. Health implications for Faroe Islanders of heavy metals and PCBs from pilot whales. *Sci Total Environ*. 1996;186:141-148.

44. Huisman M, Koopman-Esseboom C, Fidler V, et al. Perinatal exposure to polychlorinated biphenyls and dioxins and its effect on neonatal neurological development. *Early Hum Dev*. 1995;41:111-127.

45. Swedish Expert Group. Methylmercury in fish: a toxciological-epidemiological evaluation of risks. *Nord Hyg Tidskr*. 1971;suppl 4:1-333

46. World Health Organization. *Evaluation of Certain Food Additives and the Contaminants Mercury, Lead and Cadmium*. Geneva, Switzerland: World Health Organization; 1972. Technical Report Series No. 505.

47. Egeland G, Middaugh J. Balancing fish consumption benefits with mercury exposure. *Science*. 1997;278:1904-1905.

THE LANCET
32 Jamestown Road London NW1 7BY UK

Prenatal methylmercury exposure from ocean fish consumption in the Seychelles child development study

Gary J Myers
Philip W Davidson
Christopher Cox
Conrad F Shamlaye
Donna Palumbo
Elsa Cernichiari
Jean Sloane-Reeves
Gregory E Wilding
James Kost
Li-Shan Huang
Thomas W Clarkson

Reprinted from THE LANCET
17 May 2003 Vol. 361
No. 9370 Pages 1686-1692

Prenatal methylmercury exposure from ocean fish consumption in the Seychelles child development study

Gary J Myers, Philip W Davidson, Christopher Cox, Conrad F Shamlaye, Donna Palumbo, Elsa Cernichiari, Jean Sloane-Reeves, Gregory E Wilding, James Kost, Li-Shan Huang, Thomas W Clarkson

Summary

Introduction Exposure to methylmercury (MeHg) before birth can adversely affect children's neurodevelopment. The most common form of prenatal exposure is maternal fish consumption, but whether such exposure harms the fetus is unknown. We aimed to identify adverse neurodevelopmental effects in a fish-consuming population.

Methods We investigated 779 mother-infant pairs residing in the Republic of Seychelles. Mothers reported consuming fish on average 12 meals per week. Fish in Seychelles contain much the same concentrations of MeHg as commercial ocean fish elsewhere. Prenatal MeHg exposure was determined from maternal hair growing during pregnancy. We assessed neurocognitive, language, memory, motor, perceptual-motor, and behavioural functions in children at age 9 years. The association between prenatal MeHg exposure and the primary endpoints was investigated with multiple linear regression with adjustment for covariates that affect child development.

Findings Mean prenatal MeHg exposure was 6·9 parts per million (SD 4·5ppm). Only two endpoints were associated with prenatal MeHg exposure. Increased exposure was associated with decreased performance in the grooved pegboard using the non-dominant hand in males and improved scores in the hyperactivity index of the Conner's teacher rating scale. Covariates affecting child development were appropriately associated with endpoints.

Interpretation These data do not support the hypothesis that there is a neurodevelopmental risk from prenatal MeHg exposure resulting solely from ocean fish consumption.

Lancet 2003; **361:** 1686–92
See Commentary page 1667

Department of Neurology (Prof G J Myers MD, D Palumbo PhD), **Department of Pediatrics** (Prof G J Myers, Prof P W Davidson PhD, J Sloane-Reeves MS), **and Division of Epidemiology, Statistics, and Prevention** (C Cox PhD), **National Institute for Child Health and Development, National Institutes of Health, Department of Health and Human Services, Bethesda, USA; Departments of Biostatistics and Computational Biology** (C Cox, G Wilding PhD, J Kost PhD, L-S Huang PhD) **and Environmental Medicine** (E Cernichiari MS, Prof T W Clarkson PhD), **University of Rochester School of Medicine and Dentistry; Ministry of Health, Republic of Seychelles** (C F Shamlaye MBChB)

Correspondence to: Dr G J Myers, University of Rochester Medical Center, 601 Elmwood Avenue, Box 631, Rochester, NY 14620, USA (e-mail: gary_myers@urmc.rochester.edu)

Introduction

Methylmercury (MeHg) is highly and selectively toxic to the CNS. The prenatal period is believed to be the most susceptible stage of life.[1] MeHg inhibits processes fundamental to brain development such as neuronal cell division and migration.[2] Although outbreaks of MeHg poisoning have occurred, human exposure is almost always from fish and other seafood. Inorganic mercury (Hg) occurring naturally or from pollution can be converted to MeHg by micoorganisms and is bioaccumulated up the food chain.[3] Ingested MeHg is almost totally absorbed and readily crosses the placenta and blood-brain barriers. Pregnant women who consume fish expose the fetus to MeHg.

Studies[4] in Iraq raised concern that prenatal MeHg exposure at the concentrations achieved by maternal consumption of ocean fish might adversely affect a child's neurodevelopment. However, the poisoning in Iraq resulted from seed grain treated with MeHg in a subsistence desert community and its relevance to fish consumers was unclear. If consumption of fish is associated with adverse neurodevelopmental consequences from MeHg exposure, public-health measures could reduce the risk. However, fish is an important source of protein worldwide and also has health benefits for adults. Consequently, restricting its consumption might adversely affect health.[5-7]

The Seychelles Child Development Study was specifically designed to test the validity of this hypothesis in a well-nourished population exposed to MeHg only from high consumption of unpolluted ocean fish. In Seychelles, women of childbearing age consume fish containing similar concentrations of MeHg to those in the USA (average about 0·3 μg/g), but the fish consumption rate is much higher (average 12 meals per week in our cohort compared with one or fewer in the USA). Consequently, concentrations of mercury in hair and other indicator media are many times higher than those in the US population and any effects should be detectable earlier. The main cohort was established in 1989 and previous assessments when the children were 6, 19, 29, and 66 months of age have been described in detail elsewhere.[8-11] Up to now, we have detected no adverse effects. We now report the results of our investigation at 9 years of age.

Methods

Participants

In 1989–90 we enrolled 779 mother-child pairs (about 50% of live births during that period), when the children were 6 months old. We excluded mothers and children with disorders highly associated with adverse neurodevelopment such as traumatic brain injury, meningitis, epilepsy, and severe neonatal illnesses. Although these disorders have been associated with overt and subtle neurodevelopmental problems, no data exist to suggest they are associated with MeHg exposure.[9] The 44 exclusions through 66 months of age have been reported.[9-11] We subsequently excluded 18 children for

Covariate	Definition
Sex	Male or female
Examiner	1, 2, or 3
Family resource scale	Continuous
Family status code	2, 1, or no biological parents in home
HELPS*	Continuous
Child's age at testing	Continuous
Child's medical history	Positive if diagnosed intrauterine growth retardation at birth or head circumference greater than 2 SD from normal
Maternal age	Continuous
HOME score†	≤31, >31 to 35, or >35
K-Bit‡	Low (<16), normal (16–28), or high (>28)
Hollingshead socioeconomic status	Unskilled (≤19), semiskilled (>19 to 29), skilled (>29 to 39), or minor/major business/profession (>39)
Hearing (best ear)	Normal (≤25 dB), borderline (>25 to 35 dB), or abnormal (> 35 dB)
Child's mercury concentration	Continuous

*Henderson early learning process scale. †HOME=home observation for measurement of the environment. ‡K-Bit=Kaufman brief intelligence test to determine caregiver intelligence.

Table 1: **Covariate definitions**

closed head trauma and meningitis. 717 (92%) of the 779 children were still eligible at 9 years of age. Of those eligible, 74 (10%) were not tested, leaving a total of 643 children (83% of the original cohort and 90% of those still eligible). The reasons children were not tested included residing abroad, refusal, and inability to locate them. The research protocols were reviewed and approved by the Institutional Review Boards of the University of Rochester and the Republic of Seychelles.

Procedures

Prenatal exposure to MeHg was determined by measuring total Hg in maternal hair growing during pregnancy, the method most indicative of prenatal exposure.[3,12] We obtained maternal hair at delivery and enrolment and measured the Hg concentrations in the sample that best recapitulated exposure. We assumed a growth rate of 1·1 cm per month and a delay of 20 days between current blood concentrations and appearance of the Hg in the first centimetre of scalp hair. Concentrations of total Hg in the hair of fish-eating populations correlate highly with maternal blood concentrations and are a better exposure index than concentrations of organic Hg.[13] Organic Hg can be partly transformed to inorganic Hg so that the concentration of total Hg more accurately represents the MeHg entering the hair follicle from the blood stream.[13] Concentrations of total Hg in maternal hair at delivery correlated highly with concentrations of Hg in brain samples taken at autopsy from Seychellois infants who died of natural causes.[14] We measured total and inorganic Hg with cold vapour atomic absorption spectroscopy with quality control procedures.[12] Concentrations of Hg are expressed in μg/g, where 1 μg/g=1 part per million (ppm) in hair.

We assessed neurocognitive, language, memory, motor, perceptual-motor, and behavioural functions. Our tests included overall and domain-specific items, covering neurodevelopmental domains associated with prenatal MeHg exposure and included most of the specific tests used in previous studies. Individual tests measured intelligence (the Wechsler intelligence scale for children III [WISC III] full-scale IQ); learning and achievement (the Woodcock-Johnson test of achievement, letter-word recognition, and applied problems subtests and the California verbal learning test); memory (the visual memory subtest of the wide-range assessment of memory and learning); motor functions (finger tapping, trailmaking, grooved pegboard, and most of the Bruininks-Oseretsky test of motor proficiency); language (Boston naming test); visual-motor integration (the Beery-Buktenica developmental test of visual motor integration and a test of haptic matching[15]); and sustained attention (Connor's continuous performance test). We assessed behaviour with the Connor's teacher rating scale and the parent-child behaviour checklist.

A team of three Seychellois child health and development professionals (a senior nurse, a child psychologist, and a special educator) assessed the children.[16] They received extensive training in child development and psychometric assessment procedures at the University of Rochester before the assessments. All personnel working in Seychelles were unaware of the MeHg exposure from the start of the study and no individual MeHg concentrations have been shared with families, clinical investigators, or anyone in Seychelles.

We investigated test reliability among testers and between each tester and a psychologist (PWD or DP). Pair-wise intertester reliability was assessed once a week by having a child's performance during testing scored by two team members simultaneously. We investigated gold-standard reliability by on-site simultaneous scoring of about 5% of test sessions by one of the psychologists.

The cohort was initially tested between February, 1997, and November, 1998. Every child was seen twice about 1 month apart. The sessions lasted about 3 h. During the first session the caregiver completed a demographic questionnaire and the parental child behaviour checklist while the child was given the WISC III and had audiometry and tympanometry. All remaining tests were administered individually during the second session. Testing took place in a specially established child development centre, mostly in the morning, and the tests were given in the same sequence in each session.

	Normal (mean [SD] or normal range)	Overall data (mean [SD])	Data by MeHg in maternal hair (mean [SE])					Regression coefficient (SE)*	p	95% CI for a 10 μg/g change
			≤3 (n=135)	>3 to 6 (n=190)	>6 to 9 (n=143)	>9 to 12 (n=87)	>12 (n=88)			
WISC III full scale IQ	100 (15)	81·6 (11·6)	79·4 (1·0)	81·1 (0·9)	80·3 (1·0)	80·8 (1·2)	81·7 (1·2)	–0·13 (0·10)	0·20	–3·3 to 0·7
CVLT										
Short delay recall	0 (1)	–0·02 (1·04)	–0·1 (0·1)	–0·6 (0·1)	0·0 (0·1)	0·0 (0·1)	0·1 (0·1)	0·013 (0·010)	0·19	–0·1 to 0·3
Long delay recall	0 (1)	–0·1 (1·03)	–0·2 (0·1)	–0·2 (0·1)	0·0 (0·1)	0·0 (0·1)	0·0 (0·1)	0·011 (0·010)	0·28	–0·1 to 0·3
BNT total score	43 (5)	26·5 (4·8)	26·3 (0·4)	26·6 (0·4)	25·9 (0·4)	27·3 (0·5)	26·7 (0·5)	–0·012 (0·046)	0·79	–1·0 to 0·8
W-J test										
LW recognition	90–110	131·7 (40·3)	130·4 (3·6)	126·3 (3·1)	135·3 (3·4)	131·3 (4·6)	140·1 (3·8)	0·19 (0·39)	0·62	–5·8 to 9·6
Applied problems	90–110	95·4 (15·5)	94·8 (1·4)	95·1 (1·2)	94·1 (1·3)	97·2 (1·8)	97·2 (1·7)	–0·057 (0·15)	0·71	–3·5 to 2·4

Higher scores in all tests indicate better performance. WISC III=Wechsler intelligence scale for children version III. CVLT=California verbal learning test. BNT=Boston naming test. W-J=Woodcock-Johnson test of achievement. LW Recognition=letter word recognition. *Data from prenatal mercury term in the multiple regression models. †95% CI, on the scale of the original measurement.

Table 2: **Results of neurodevelopment tests for cognition and achievement by prenatal exposure**

	Normal (mean [SD] or normal range)	Overall data (mean [SD])	Data by MeHg in maternal hair (mean [SE])					Regression coefficient (SE)*	p	CI for a 10 μg/g change†
			≤3 (n=135)	>3 to 6 (n=190)	>6 to 9 (n=143)	>9 to 12 (n=87)	>12 (n=88)			
VMI	100 (15)	96·0 (11·7)	95·1 (1·2)	95·8 (0·8)	96·7 (0·9)	95·7 (1·2)	96 6 (1·4)	–0·010 (0·12)	0·93	–2·4 to 2·2
Bruininks-Oseretsky‡	··	44·6 (6·1)	44 (0·6)	44·3 (0·5)	45 (0·5)	45 (0·6)	45·3 (0·6)	0·093 (0·056)	0·10	–0·2 to 2·0
Haptic discrimination test (total correct out of 10)‡	··	4·1 (1·8)	4·2 (0·2)	4·1 (0·1)	4 (0·2)	4·1 (0·2)	4·2 (0·2)	–0·010 (0·018)	0·60	–0·5 to 0·3
Grooved pegboard time (s)§										
Dominant hand¶	74 (15)	91·8 (20·5)	86 (78–100)	88·5 (79–100)	89 (80–100)	87 (78–101)	89 (78–99)	$3·3\times10^{-5}$ ($1·9\times10^{-5}$)	0·08§	91·4 to 98·1
Non-dominant hand¶	81 (16)									
Male		100·1 (18·9)	95 (87–104)	95·5 (86–113)	93 (84–105)	106 (93–113)	98 (89–110)	$6·5\times10^{-6}$ ($2·5\times10^{-5}$)	0·01	101·7 to 112·9
Female		108·2 (29·8)	106 (95–124)	101 (90–113)	107·5 (89–120)	101 (85–120)	98 (90–113)	$-2·5\times10^{-6}$ ($2·6\times10^{-5}$)	0·34	100·0 to 111·3
Trail making time (s)										
A§	25 (9)	33·7 (17·1)	30 (24–40)	29 (22–40)	29 (23–39)	31 (25–39)	31 (23–38)	0·0037 (0·0038)	0·33	32·5 to 37·6
B§	55 (19)	81·5 (49·6)	67 (49–101)	63 (48–90)	65 (52–101)	63·5 (49–91)	62 (47–89)	0·0067 (0·0050)	0·17	79·1 to 96·0
Finger tapping										
Dominant hand	M 40 (5) F 39 (5)	34·0 (5·7)	34 (0·5)	34 (0·4)	33·8 (0·5)	34·7 (0·6)	33·7 (0·5)	–0·050 (0·053)	0·34	–1·5 to 0·5
Non-dominant hand	M 35 (5) F 33 (5)	30·0 (4·6)	29·5 (0·4)	30 (0·4)	30·1 (0·4)	31 (0·5)	29·8 (0·4)	0·016 (0·041)	0·69	–0·6 to 1·0
WRAML design memory	10 (3)	7·7 (2·9)	7·7 (0·2)	7·6 (0·2)	7·9 (0·3)	7·7 (0·3)	7·8 (0·3)	–0·021 (0·029)	0·48	–0·8 to 0·4

Higher scores indicate better performance except for Grooved Pegboard and Trailmaking where higher scores indicate poorer performance. VMI=visual motor integration. WRAML=wide range assessment of memory and learning. *Data from prenatal mercury term in the multiple regression models. †95% CI on the scale of the original measurement. ‡No norm available for the Bruininks-Oseretsky since the complete test was not administered. §The prenatal MeHg by sex main interaction was significant. ¶Where indicated test score was transformed for analysis. Medians and quartiles are reported for untransformed scores. Confidence limits should be compared with the mean test score (third column); if the slope was non-significant the confidence limits straddle this value.

Table 3: **Neurodevelopmental tests for motor, perceptual motor and memory by prenatal exposure**

When the cohort was 42–56 months of age, a home visit was made to administer the Caldwell-Bradley preschool version of the home observation for measurement of the environment. The child's primary caregiver, defined as the family member with whom the child lived for at least 5 days per week hosted the visit (93% of caregivers were the child's biological mother). Primary caregivers were recalled in 1999–2000 and given the family resource scale and the Henderson early learning process scale to measure the quality of stimulation in the home environment. They also completed the matrices subtest of the Kaufman brief intelligence test to determine caregiver intelligence.

The covariates used in the analysis are shown in table 1. They were selected a priori for their known effect on child development and were expected to provide an index of the effectiveness of the assessments. The Hollingshead four-factor socioeconomic score was calculated with a list of Seychellois employment codes. Recent postnatal MeHg exposure was included since it was associated with outcomes in the 66-month assessments. It was measured in a 1-cm segment of hair closest to the scalp on a sample taken at the initial 9-year assessment, and represented about 1 month of recent exposure. The mean postnatal hair concentration was 6·1 μg/g (SD 3·5). Lead, polychlorinated biphenyls, and pesticides were not

	Normal (mean [SD] or normal range)	Overall data (mean [SD])	Data by MeHg in maternal hair (mean [SE])					Regression coefficient (SE)*	p	CI for a 10 μg/g change†
			≤3 (n=135)	>3 to 6 (n=190)	>6 to 9 (n=143)	>9 to 12 (n=87)	>12 (n=88)			
CPT										
Hit reaction time‡	50 (10)	31·6 (14·1)	29·2 (1·3)	33·6 (1·1)	32·2 (1·2)	31·7 (1·7)	29·6 (1·9)	–0·13 (0·16)	0·41	(–4·4 to 1·8)
Attentiveness	50 (10)	57·5 (9·7)	57·3 (0·9)	58·8 (0·8)	56·7 (0·9)	57·2 (1·3)	56·8 (1·1)	–0·0063 (0·10)	0·95	(–2·1 to 2·0)
Risk-taking‡	50 (10)	75·0 (29·7)	74·3 (2·0)	75 (1·7)	77·6 (4·4)	71·7 (2·3)	75·3 (2·5)	0·11 (0·22)	0·60	(–3·1 to 5·4)
CBCL	50 (10)	59·4 (10·2)	59·8 (0·9)	59·4 (0·8)	58·7 (0·9)	58·3 (1·1)	60·9 (1·1)	–0·031 (0·10)	0·76	(–2·3 to 1·7)
CTRS hyperactivity index§	45–55	55·3 (12·8)	52 (44–65)	51·5 (46–65)	52 (45–60)	50 (46–59)	49 (44–60)	–0·0067 (0·0023)	0·004	(49·4 to 54·1)

Higher scores indicate better performance except for the CBCL and TRS where higher scores indicate poorer performance. CPT=continuous performance task. CBCL=Connor's child behaviour checklist. CTRS=Connor's teacher rating scale. *Data from prenatal mercury term in the multiple regression models. †95% CI on the scale of the original measurement. ‡A score that deviates positively or negatively from the norm can be regarded clinically indicative of attention problems. §Where indicated test score was transformed for analysis. Confidence limits should be compared with the mean test score (third column); if the slope was non-significant the confidence limits straddle this value. ¶Medians and quartiles are reported for untransformed scores.

Table 4: **Neurodevelopmental tests for attention and behaviour by prenatal exposure**

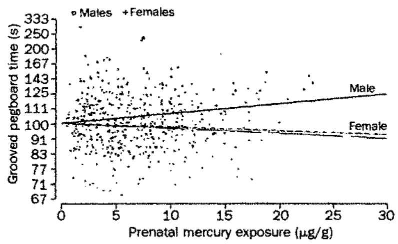


Figure 1: **Association between prenatal MeHg exposure and scores on the groove pegboard with the non-preferred hand**
Test scores have been adjusted for covariates to show the association with mercury concentrations in the multiple regression model. Lines are shown for the model with outliers (dashed) and without outliers (solid). Outliers are indicated by larger symbols. Test scores were transformed for analysis so the y-axis is nonlinear.

included since measured concentrations in Seychelles are low.[10,17] Use of alcohol and tobacco among Seychellois women of childbearing age is low; consequently they were not used as covariates. At enrolment, 5% of the mothers reported occasional alcohol intake during pregnancy and 2% reported tobacco use.

Statistical analysis

The primary analysis included 21 endpoints. We used the main score for most tests, but several measures yielded more than one endpoint. For every endpoint, we did a maximum of three linear-regression analyses for prenatal MeHg exposure using all the covariates defined in table 1. All analyses were done with the SAS system, version 8. Because differential effects on males and females have been reported, every model was run first with and then without a MeHg by sex interaction term for both prenatal and recent postnatal exposure.[10,18–20] If the overall test for both models was not significant at a two-tailed significance level of 0·05, the results of that analysis were deemed negative. We assessed the model with interactions first. If both interactions were significant then results are reported for this model. If neither interaction was significant, we report the model without interactions. If only the prenatal or postnatal interaction were significant, we reran the model, dropping the non-significant

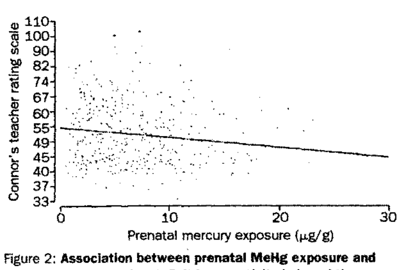


Figure 2: **Association between prenatal MeHg exposure and scores on the attention deficit hyperactivity index of the Connor's teacher rating scale**
Test scores have been adjusted for covariates to show the association with mercury concentration in the multiple regression model. Lines are shown for the model with (dashed) and without (solid) outliers. Outliers are indicated by black ovals. Test scores were transformed for analysis so the y-axis is non-linear.

interaction term (including only a simple effect), but including the significant interaction term. This happened for four endpoints.

Every analysis included an assessment of residuals as a check on the assumptions of normally distributed errors with constant variance. If the assumptions seemed to be violated, we used power transformations to stabilise the variance and produce more normally distributed errors. For every analysis, we assessed the model for statistical outliers (scores with standardised residual values >3 or <–3). All models with outliers were rerun without the outliers and the results with and without outliers were compared and are reported.

We also assessed every regression model for the effects of influential points, identified by deleting each point in the data set individually from the analysis and calculating the resulting standardised change in the regression coefficient for prenatal MeHg exposure. The regression analysis for all primary endpoints was repeated without influential points to determine whether the original results were dependent upon such points. The final analysis included influential points that were not also outliers.

Role of the funding source

The sponsors of the study approved the study design but had no other involvement in the study design, data collection, data analysis, data interpretation, or writing of this report.

Results

The mean prenatal total MeHg exposure was 6·9 µg/g (SD 4·5). The correlation coefficient between prenatal and postnatal exposure was –0·08 (p=0·04).

Intraclass correlation coefficients were computed for each of the 13 subtests of the WISC III. This test was chosen for reliability computations since it was the most difficult component to administer. Comparisons between WISC-III subtest scores obtained by pairs of testers (n=45) ranged from 0·90 to 1·00. Agreement between each tester and one of the team psychologists (n=37) ranged from 0·81 to 1·00.

The mean age at testing was 107 months (SD 4). For data presentation, the mean (SD) for each endpoint was computed by prenatal mercury exposure groupings. These results are shown in table 2 (cognition and achievement domains), table 3 (motor, perceptual-motor, and memory domains), and table 4 (attention and behaviour domains). In cases where the endpoint was transformed for analysis, the median and quartiles have been included in the tables, rather than the mean. Seychellois children's scores on

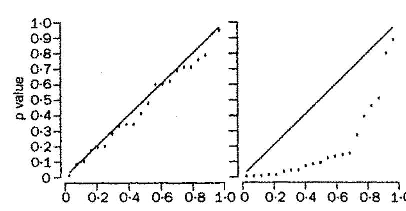


Figure 3: **Probability plots of p values for the prenatal exposure effect (left) and the home observation covariate (right) from 21 primary regression analyses**
The degree to which p values conform to the null hypothesis line shows the consistency of the data with the overall null hypothesis.

most endpoints compared favourably with US norms. The variability of the tests was as expected and they seemed to discriminate well among cohort children. The WISC III and the Boston naming test (table 2) were both affected by cultural variation, with lower means for Seychellois children than for US controls. However, the variability associated with these and other endpoints was consistent with test expectations, suggesting that all tests discriminated well among cohort children. Seychellois children did substantially better than the US norms on the Woodcock-Johnson letter-word recognition test, measuring scholastic achievement in reading.

Significant two-tailed overall model *F* statistics ($p<0{\cdot}05$) resulted from the regression analysis for 20 of the 21 models with outliers removed. The squared multiple correlation coefficients for these models ranged between 5% and 24% with most values between 10 and 20%. Significant associations between prenatal MeHg exposure and performance were found for two endpoints, and both needed transformation for analysis. There was a significant decrease in performance on the grooved pegboard time for the non-dominant hand in males (table 3; figure 1). This task required an average of 100 s for cohort participants to complete and a 95% CI for a 10 μg/g change in exposure ranged from 1 to 13 s. The variability for this task within the reference population was 19 s, suggesting the effect was small. The association for the same task using the dominant hand was not significant. There was a significant improvement of the hyperactivity index of the Connor's teacher-rating scale (table 4; figure 2) as prenatal MeHg increased. A 10 μg/g increase in exposure would result in a drop of between 1 and 6 points (95% CI) in the hyperactivity index. Seychellois children were functioning at the upper limit of the normal range for this test.

	Caregiver IQ	Maternal	Child's age	Home: Overall	Home: Low	Home: Med	Home: High	Socioeconomic score: Overall	Unskill	Semi-skill	Skill	Professional
WISC III full scale IQ	0·14	0·20	−2·08		−1·61	−0·40	2·02		−2·61	0·28	−0·20	2·53
	(0·03)*	(0·08)†	(1·48)	*	(0·67)	(0·64)	(0·68)	*	(0·8)	(0·75)	(0·84)	(1·06)
CVLT												
Short delay	0·001	−0·01	−0·61		−0·11	0·05	0·06		−0·17	−0·0004	−0·05	0·22
	(0·003)	(0·01)	(0·16)*		(0·07)	(0·06)	(0·07)	†	(0·08)	(0·08)	(0·09)	(0·11)
Long delay	0·0001	−0·002	−0·57		−0·05	−0·04	0·09		−0·26	0·03	0·10	0·13
	(0·003)	(0·01)	(0·16)*		(0·07)	(0·06)	(0·07)	†	(0·08)	(0·08)	(0·09)	(0·11)
Boston naming test total score	0·03	0·06	1·45		−1·10	0·13	0·97		−0·94	−0·34	0·22	1·06
	(0·02)‡	(0·04)‡	(0·73)†	*	(0·31)	(0·29)	(0·31)	*	(0·36)	(0·34)	(0·38)	(0·48)
W-J test												
Letter-word recognition	0·44	0·64	13·7		1·37	−7·60	6·23		−3·69	−0·34	1·08	2·95
	(0·13)*	(0·32)†	(8·03)	*	(2·73)	(2·52)	(2·7)		(3·18)	(2·94)	(3·29)	(4·14)
Applied problems	0·24	0·20	0·96		−1·21	−0·79	2·00		−3·62	0·66	−1·03	3·98
	(0·05)*	(0·12)*	(3·00)		(1·02)	(0·93)	(1·02)	*	(1·17)	(1·11)	(1·24)	(1·56)
VMI	0·11	0·03	−3·99		−1·03	−0·68	1·71		−1·33	0·98	0·86	−0·52
	(0·04)*	(0·09)	(1·86)†		(0·78)	(0·73)	(0·78)		(0·91)	(0·86)	(0·96)	(1·2)
B-0 test of motor development	0·03	−0·01	2·42		−0·004	−0·36	0·36		−0·14	−0·43	0·32	0 26
	(0·02)‡	(−0·04)	(0·88)*		(0·38)	(0·35)	(0·38)		(0·44)	(0·42)	(0·46)	(0·58)
Haptic discrimination test	0·01	−0·02	−0·37		−0·03	0·06	−0·03		0·01	0·01	−0·08	0·07
	(0·01)	(0·01)	(0·29)		(0·12)	(0·12)	(0·12)		(0·14)	(0·14)	(0·15)	(0·19)
Grooved pegboard												
Dominant hand§	−1·9	−3·7	−92		28	4·5	−32		1·3	20	20	2·7
	(0·63)	(1·5)	(31)		(13)	(0·12)	(13)	†	(15)	(14)	(16)	(20)
Non-dominant hand§	1·4	−02·1	−100		36	−1·0	−35		15	4·6	−2·8	−17
	(0·6)	(1·4)	(29)		(12)	(11)	(12)	*	(14)	(13)	(15)	(19)
Trail-making												
A	−0·003	−0·004	−0·30		0·03	−0·02	−0·01		0·0001	0·01	0·01	−0·02
	(0·001)†	(0·003)	(0·06)*		(0·03)	(0·02)	(0·03)		(0·03)	(0·03)	(0·03)	(0·04)
B	−0·01	−0·01	−0·35		0·05	0·04	−0·09		0·04	0·05	0·03	−0·12
	(0·002)*	(0·004)	(0·08)*	*	(0·03)	(0·03)	(0·03)		(0·04)	(0·04)	(0·04)	(0·05)
Finger tapping												
Dominant hand	0·02	0·06	3·33		−0·76	0·09	0·67		−0·31	−0·30	0·45	0·16
	(0·02)	(0·04)	(0·85)*		(0·36)	(0·33)	(0·36)		(0·42)	(0·39)	(0·44)	(0·56)
Non-dominant hand	0·003	0·05	2·05		−0·98	0·28	0·70		−0·27	−0·42	0·40	0·29
	(0·01)	(0·03)	(0·67)*	*	(0·28)	(0·26)	(0·28)		(0·33)	(0·31)	(0·35)	(0·44)
WRAML design memory	0·03	0·02	−0·22		−0·43	0·05	0·38		−0·25	−0·03	0·08	0·19
	(0·01)*	(0·02)	(0·47)		(0·20)	(0·19)	(0·20)		(0·23)	(0·22)	(0·24)	(0·31)
CPT												
Hit reaction	0·13	0·20	0·99		−2·61	0·92	1·69		0·33	0·10	−0·56	0·13
	(0·05)†	(0·12)‡	(2·12)	†	(1·03)	(0·96)	(1·03)		(1·22)	(1·14)	(1·26)	(1·57)
Attentiveness	−0·05	−0·17	3·87		1·31	−0·37	−0·93		0·003	0·95	−0·63	−0·32
	(0·03)	(0·08)†	(1·39)*		(0·68)	(0·64)	(0·68)		(0·8)	(0·75)	(0·83)	1·03
Risk-taking	−0·16	0·26	3·18		2·84	−0·75	−2·09		0·12	0·39	0·46	−0·97
	(0·07)†	(0·16)	(2·9)		(1·14)	(1·32)	(1·42)		(1·67)	(1·55)	(1·72)	(2·14)
CBCL	−0·06	−0·32	1·58		0·70	0·68	−1·38		−0·21	−0·44	−0·13	0·77
	(0·03)	(0·08)*	(1·48)		0·68	(0·64)	(0·69)		(0·79)	(0·76)	(0·84)	(1·06)
CTRS hyperactivity index	−0·001	0·0003	−0·05		−0·01	0·01	0·004		0·02	0·02	−0·01	−0·02
	(0·001)	(0·002)	(0·03)‡		(0·02)	(0·01)	(0·02)		(0·02)	(0·02)	(0·02)	(0·02)

Values are regression coefficients (SE). WISC=Wechsler intelligence scale for children. CVLT=California verbal learning test. W-J=Woodcock-Johnson test of achievement. VMI=visual motor integration. B-0=Bruininks-Oseretsky. WRAML=wide range assessment of memory and learning. CPT=continuous performance test. CBCL=children's behaviour checklist. CTRS=Connors Teacher Rating Scale. *$p<0{\cdot}01$. †$p<0{\cdot}05$. ‡$p<0{\cdot}10$. †Where indicated test score was transformed for analysis. §Values are $\times 10^{-5}$.

Table 5: **Social and environmental covariate effects**

The large number of endpoints raised concern about multiplicity and so we investigated the actual distributions of the p values for the prenatal exposure for consistency with the overall null hypothesis of no association. If the overall null hypothesis were true, the p values should be uniformly distributed from zero to one. We assessed this graphically by plotting the p values against idealised values from the uniform distribution. The distribution of p values was consistent with the expected values, and provides support for the overall null hypothesis of no association (figure 3). For comparison we plotted the p values for the home observation for measurement of the environment, a covariate from the same series of analyses that was associated with the endpoints.

Table 5 shows the associations between endpoints and selected social and environmental covariates. These factors have well established associations with child development and were expected to provide an index of the effectiveness of the assessments in ascertaining developmental status of the children. The data suggest that the effects of these factors on the endpoints were consistent with established associations. For example, socioeconomic score, early home environment scores, and maternal IQ were consistently associated with outcomes for neurocognitive endpoints but only occasionally with outcomes on motor tasks. Measures of later home environment, such as the family resource scale and the Henderson early learning process scale had limited effect, as would be expected for normally developing children at this age, and are not included.

We have focused on prenatal exposures, with postnatal hair concentration of MeHg a covariate in the analysis for prenatal effects. In a few tests this analysis suggested an adverse association with postnatal exposure in females. Since postnatal exposure differs substantially from prenatal exposure and since males are thought to be more susceptible, the interpretation of these findings is unclear. Analyses for the entire postnatal period are still in progress.

19 regression analyses revealed between one and five outlier scores involving a total of 40 different participants. In all cases, the association between prenatal MeHg exposure and the endpoint was the same, irrespective of whether outliers were included. For the two endpoints with a significant prenatal MeHg effect (Connor's teacher rating scale [two outliers] and the grooved pegboard non-dominant hand [three outliers]), the outlier scores all had prenatal MeHg concentrations of 7·5 μg/g or less and low performance.

Every model had between 0 and 3 influential points, defined as a score that may have affected the slope of the regression line but did not reach the status of an outlier. We report results with influential points included. However, the models with both interactions were re-run without influential points included, and in no case did the results for prenatal exposure change.

Discussion

Two of 21 endpoints were associated with prenatal MeHg exposure and developmental outcomes at 9 years of age. One association involved diminished performance (grooved pegboard non-dominant hand in males only) and the other an enhancement (hyperactivity index of the Connors teacher rating scale). As indicated by the distribution of p values in figure 3, both these outcomes are probably due to chance.

Results of studies of prenatal exposure to MeHg from seafood consumption in the Faeroe Islands[21] and New Zealand[22,23] have shown adverse neuropsychological outcomes in school-aged children. The difference in findings from these studies and the Seychelles study has been investigated in two reviews[6,7] and several explanations have been proposed,[24] including the power of the studies to detect subtle differences. Assessment of standard power curves for the two studies shows that the power of the Seychelles study was only slightly less than that of the Faeroe Islands and both are substantially greater than any previous study. Our original power calculations estimated a 90% chance of detecting a five-point difference on the Bayley scales of infant development with every 10 μg/g increase in MeHg.

These outcomes might differ because of the cellular effect of very different concentrations of MeHg in the seafood consumed by these populations. The presence of cellular mechanisms in mammals that detoxify Hg raises the possibility that a larger bolus dose with a meal might behave differently than a small dose.[25] In Seychelles, the seafood consumed has a lower concentrations of MeHg than in the other two populations. The mean concentration of organic Hg in whale meat (the main source of MeHg in the Faeroe Islands) was 1·6 μg/g (SD 0·4).[26] In New Zealand the shark muscle consumed in the popular take-out food of fish and chips had a mean Hg concentration of 2·2 μg/g with some samples more than 4 μg/g.[27] By contrast, the Seychellois consume many different species of ocean fish, the mean MeHg content of which averages 0·3 μg/g (with 97·5 % of the samples below 0·7 μg/g).[12]

Cord blood was used as the monitoring medium in the Faeroe Islands and the investigators argued that it is a more sensitive biomarker for prenatal MeHg exposure than concentrations in hair.[21] However, maternal hair has been the biological monitor of choice in most studies of prenatal exposure and was used in the New Zealand study.[3] Moreover, hair and blood concentrations are closely correlated, and hair can recapitulate exposure during the entire period of pregnancy.[12]

The tests used and the age at testing also differed between the studies.[6] However, our test battery included both global and domain-specific items and nearly all the tests reported previously had shown an association with MeHg.[21,23] Moreover, most of the tests gave results that were normally distributed and scores that were similar to those in Western countries, and were sensitive enough to detect the expected effects of covariates. The tests should have been sensitive enough to detect MeHg effects if they were present. A difference in age at testing is no longer a viable explanation since our previous findings at 66 months of age are now extended to 9 years, thus bracketing the 6 and 7 years ages used in the other studies.

One factor unique to the Faeroe Islands study is the consumption of whale meat and blubber. Whale blubber has high concentrations of polychlorinated biphenyls and other persistent organic pollutants and the meat has concentrations of inorganic Hg similar to MeHg.[24,26] These factors would not explain the associations reported from New Zealand.

Our study has potential limitations. As with any observational study, enrolment could have been biased. We sought to prevent such bias by offering participation to all children who reached the target age, and comparisons of enrolled children with those not participating did not show any bias. The best biomarker for prenatal exposure to MeHg is also controversial. In studies of neurodevelopmental outcomes, choosing the appropriate covariates affecting child development is crucial to get valid results from the statistical models. We included a standard

array of covariates known to be related to developmental outcomes that were selected a priori. A-priori selection was preferred to alternative data-driven approaches to avoid the possibility of bias and type-1 errors. Because children were enrolled when they were 6 months old, information about pregnancy, birth, and feeding was obtained retrospectively and might not have been as accurate as if it had been gathered at delivery. We also did not assess nutritional covariates such as selenium or polyunsaturated fatty acids. We are now investigating a new cohort in Seychelles to address this issue. Different results might have been obtained with different developmental tests, but we designed our tests to assess developmental domains known to be affected by prenatal MeHg poisoning and included most tests administered in previous studies. Finally, the choice of a different statistical approach might have led to different results. However, secondary analyses of our earlier studies with other statistical models have been consistent with conclusions based on the primary analysis.[29,30]

In summary, the Seychelles Child Development Study longitudinal assessments at 9 years of age indicate no detectable adverse effects in a population consuming large quantities of a wide variety of ocean fish. These results are consistent with our earlier findings in the same children when examined at 6, 19, 29, and 66 months of age. In Seychelles, fetal exposure was continuous through frequent consumption of ocean fish containing concentrations of MeHg comparable to those consumed by the general population in the USA. We recorded effects from covariates known to affect child development, but did not find an association with prenatal mercury. We believe this finding is relevant to public health measures and that the Seychelles could serve as a sentinel population for fish consumers.

Contributors

G Myers, P Davidson, C Cox, T Clarkson, C Shamlaye, and D Palumbo designed and helped inplement the study. C Shamlaye and G Myers managed the study in Seychelles and Rochester, respectively. P Davidson and D Palumbo trained testers and did on site reliabilities. E Cernichiari analysed mercury and managed the data office. J Sloane-Reeves scored tests, trained testers, and with E Cernichiari managed the database. C Cox, G Wilding, J Kost, and L-S Huang were responsible for statistical analysis. G Myers, P Davidson, C Cox, and T Clarkson interpreted the data and wrote the report.

Conflict of interest statement

None declared.

Acknowledgments

This research was supported by Grants R01-ES10219; R01-08442; ES-01247, and T32 ES-007271 from the US National Institutes of Health; the Food and Drug Administration, USDHHS, and by the Ministry of Health, Republic of Seychelles.

References

1 Harada Y. Congenital (or fetal) Minamata disease. In: Study Group of Minamata Disease eds. Minamata Disease. Japan; Kumamoto University, 1968: 93–118.
2 Choi BH, Lapham LW, Amin-Zaki L, Saleem T. Abnormal neuronal migration, deranged cerebral cortical organization and diffuse white matter astrocytosis of human fetal brain: a major effect of methylmercury poisoning in utero. *J Neuropathol Exp Neurol* 1978; **87:** 719–33.
3 WHO. Environmental Health Criteria 101 Methylmercury. Geneva: World Health Organization, 1990.
4 Cox C, Clarkson TW, Marsh DO, Amin-Zaki L, Tikriti S, Myers GJ. Dose-response analysis of infants prenatally exposed to methylmercury: an application of a single compartment model to single-strand hair analysis. *Environ Res* 1989; 31: 640–49.
5 Daviglus ML, Stamler J, Orenica AJ, et al. Fish consumption and the 30-year risk of fatal myocardial infarction. *N Engl J Med* 1997; **336:** 1046–53.
6 National Institute of Environmental Health Sciences. Workshop report on scientific issues relevant to assessment of health effects from exposure to methylmercury. (Raleigh, November 18–20, 1998). http://ntp-server.niehs.nih.gov/main_pages/PUBS/MethMercWkshpRpt.html#MeHgTOC (accessed Oct 8, 2001).
7 National Research Council. Toxicological effects of methylmercury. Washington, DC: National Academy Press, 2000: 1–344.
8 Marsh DO, Clarkson TW, Myers GJ, et al. The Seychelles study of fetal methylmercury exposure and child development: Introduction. *Neurotoxicology* 1995; **16:** 583–96.
9 Myers GJ, Marsh DO, Davidson PW, et al. Main neurodevelopmental study of Seychellois children following in utero exposure to methylmercury from a maternal fish diet: Outcome at six months. *Neurotoxicology* 1995; **16:** 653–64.
10 Davidson PW, Myers GJ, Cox C, et al. Effects of prenatal and postnatal methylmercury exposure from fish consumption on neurodevelopment: outcomes at 66 months of age in the Seychelles Child Development Study. *JAMA* 1998; **280:** 701–07.
11 Davidson PW, Myers GJ, Cox C, et al. Longitudinal neurodevelopment study of Seyuchellois children following in utero exposure to methylmercury from fish ingestion: outcomes at 19 and 29 months. *Neurotoxicology* 1995; **16:** 677–88.
12 Cernichiari E, Toribara TY, Liang L, et al. The biological monitoring of mercury in the Seychelles Study. *Neurotoxicology* 1995; **16:** 613–28.
13 Phelps RW, Clarkson TW, Kershaw TG, Wheatley B. Interrelationships of blood and hair mercury concentrations in a North American population exposed to methylmercury. *Arch Environ Health* 1980; **35:** 161–68.
14 Lapham LW. Cernichiari E, Cox C, et al. An analysis of autopsy brain tissue from infants prenatally exposed to methylmercury. *Neurotoxicology* 1995; **16:** 689–704.
15 Davidson P, Abbott S, Gershenfeld J. Influence of exploration time on haptic and visual matching of complex shape. *Perception Psychophysics* 1974; **14:** 539–43.
16 Davidson PW, Palumbo D, Myers GJ, et al. Neurodevelopmental outcomes of Seychellois children from the pilot cohort at 108 months following prenatal exposure to methylmercury from a maternal fish diet. *Environ Res* 2000; **84:** 1–11.
17 Davidson PW, Myers GJ, Cox C, Cernichiari E, Clarkson TW, Shamlaye C. In reply (letter to the editor). *JAMA* 1999; **281:** 897.
18 Marsh DO, Clarkson TW, Cox C, Myers GJ, Amin-Zaki L, Al-Tikriti S. Fetal methylmercury poisoning. *Arch Neurol* 1987; **44:** 1017–22.
19 McKeown-Eyssen GE, Ruedy J, Neims A. Methyl mercury exposure in northern Quebec II. Neurologic findings in children. *Am J Epidemiol* 1983; **118:** 470–79.
20 Sager PR, Aschner N, Rodier PM. Persistent differential alterations in developing cerebellar cortex of male and female mice after methylmercury exposure. *Dev Brain Res* 1984; **12:** 1–11.
21 Grandjean P, Weihe P, White RF, et al. Cognitive deficit in 7-year-old children with prenatal exposure to methylmercury. *Neurotoxicol Teratol* 1997; **19:** 417–28.
22 Kjellstrom T, Kennedy P, Wallis S, Mantell C. Physical and mental development of children with prenatal exposure to mercury from fish. Stage 1: preliminary tests at age 4. Solna, Sweden: National Swedish Environmental Board Report No 3080: 1986.
23 Kjellstrom T, Kennedy P, Wallis S, et al. Physical and mental development of children with prenatal exposure to mercury from fish. Stage 2: interviews and psychological tests at age 6. Solna, Sweden: National Swedish Environmental board Report 3642; 1989.
24 Dourson ML, Wullenweber AE, Poirier. Uncertainties in the reference dose for methlmercury. *Neurotoxicology* 2001, **22:** 677–89.
25 Clarkson TW. The three modern faces of mercury. *Environ Health Perspect* 2002; **110** (suppl): 11–23.
26 Julshamin K, Andersen A, Ringdal O, Morkore J. Trace elements intake in the Faroe Islands 1: element levels in edible parts of pilot whales (*Globicephalus meleanus*). *Sci Total Environ* 1987; **65:** 53–62.
27 Mitchell JW, Kjellstrom TE, Reeves RL. Mercury in takeaway fish in New Zealand. *N Z Med J* 1982; **95:** 112–14.
28 Cernichiari E, Brewer R, Myers GJ, et al. Monitoring methylmercury during pregnancy: maternal hair predicts fetal brain exposure. *Neurotoxicology* 1995; **16:** 705–10.
29 Axtell CD, Myers GJ, Davidson PW, et al. Semiparametric modeling of age at achieving developmental milestones after prenatal exposure to methylmercury in the Seychelles Child Development Study. *Environ Health Perspect* 1998; **106:** 559–64.
30 Axtell CD, Cox C, Myers GJ, et al, The association between methylmercury exposure from fish consumption and child development at five and a half-years of age in the Seychelles Child Development Study: an evaluation of nonlinear relationships. *Environ Res Sec A* 2000; **84:** 12–19.

The Atlanta Journal-Constitution

JUNE 6, 2003 **A12**

'Clear Skies' mercury curb put in doubt

GOP seeks relaxation of pollution cap

By JEFF NESMITH
Cox Washington Bureau

Washington – President Bush's proposal for reducing mercury emissions from electric power plants was based on inaccurate estimates, a White House aide testified Thursday. The admission prompted some Republican senators to say the plan's cap on mercury pollution should be relaxed.

"I am concerned that the cap is too stringent and creates too much uncertainty," said Sen. James Inhofe (R-Okla.), chairman of the Senate Environment and Public Works Committee.

wever, the senators were unable to ... down the White House witness, who was testifying before the air pollution subcommittee, on whether President Bush would support a relaxation of the caps.

When asked if the administration wants to change the caps, Randall Kroszner of the White House Council of Economic Advisers responded: "Occasionally, Congress does change things the president proposes, and that could be a possibility." But, he added, "We support the president's plan."

The administration's "Clear Skies" plan had projected that if pollution by sulfur and nitrogen compounds is reduced by 70 percent in 2010 as envisioned in the plan, it would automatically reduce mercury pollution from power plants to roughly 26 tons a year nationally. Today, the plants emit about 48 tons of the poisonous metal.

However, new analysis now indicates this "co-benefit" reduction in mercury pollution would amount only to between 2 and 14 tons a year, Kroszner testified.

The news was greeted with enthusiasm by power company lobbyists.

That is because the bill in Congress to implement Clear Skies sets a national mercury reduction goal of 22 tons — the amount of the original estimate of co-benefit reduction.

The companies have argued that that target would force them to spend millions of dollars on mercury control equipment. Mercury reduction goals in the bill should be trimmed back to comply with the new estimate, power company officials argue.

Environmentalists were less enthusiastic.

"This gives anti-environmentalists in Congress a weapon to weaken an already bad bill," said Frank O'Donnell, director of the Clean Air Trust.

Testifying Thursday, Southern Co. official Larry S. Monroe said that technology for removing mercury from power plant smoke is "very much in its infancy."

"There are currently no commercial technologies that are available for controlling mercury from coal-fired power plants," said Monroe, Southern program manager for pollution control research.

However, a representative of the Newark, Del., company W.L. Gore & Associates, said his company had developed a system that would capture more than 90 percent of mercury in the power plant emissions. Richard Bucher said the company will have trouble selling the new pollution technology if the government does not require power companies to reduce mercury emissions.

The National Academy of Sciences has warned that mercury can cause brain damage to the children of pregnant women who eat large quantities of fish in which the poisonous metal has accumulated. Many states, including Georgia, have issued advisories recommending that people restrict the quantity of fish consumed from lakes and rivers polluted with mercury that falls in rain.

The Environmental Protection Agency is already preparing new rules to limit the pollutant. Clear Skies would replace those rules. The White House plan would set national limits on nitrogen oxides, which cause smog, and sulfur dioxide, leading to acid rain.

The technology that reduces these two pollutants also is expected to have some effect on mercury. Therefore, the White House had estimated a co-benefit reduction of 22 tons aer year in mercury emissions.

In his testimony, Kroszner told Sen. George Voinovich (R-Ohio) that new calculations had caused EPA and the Energy Department to conclude that the incidental mercury reduction would be much less. EPA now expects Clear Skies to remove only 14 tons of mercury a year by 2010, and the Energy Department's Energy Information Administration calculates the reduction will be only 2 tons.

March 7, 2003

The Philadelphia Inquirer

Mercury rising

Bush's proposal would make a risky situation worse. Congress can stop it.

You might want to skip seconds at the fish fry.

Certain freshwater fish caught in Pennsylvania and New Jersey aren't safe to eat in large quantities. Neither are ocean swordfish, shark, king mackerel, golden snapper and, maybe, tuna.

That's because fish flesh contains mercury, which can cause palsies, seizures, learning problems, and structural abnormalities in children, even at low doses.

America needs to eliminate this toxin from its food chain.

Unlike other pollutants, mercury has only recently come under federal regulation. Forty-four states, including New Jersey in January, have had to issue warnings against fish consumption, particularly for children and pregnant women. Pennsylvania ranks seventh nationally in mercury pollution.

More than a third of mercury invades fish by falling into the water as pollution from coal-fired power plants falls from the sky. Congress and the U.S. Environmental Protection Agency have been debating since 1998 what to do about that pollution. Just as the process is about to yield rules that could reduce mercury and 58 other toxins up to 90 percent by 2008, the Bush administration has a different idea.

The President's Clear Skies Initiative - the proposed replacement for the landmark 1970 Clean Air Act - was introduced in the House and Senate last week. It would cut mercury emissions by 69 percent by 2018.

So the choice is: Eliminate more mercury sooner with the impending rules, or less, later with Bush's industry-backed plan. That shouldn't be a hard choice, Congress.

The President's program naively relies on the invisible hand of the market to reduce mercury, smog-inducing nitrogen oxide, and acid-rain-producing sulfur dioxide. Under national caps, power plants would buy and trade pollution rights. A dirty plant could buy credits from a cleaner plant, rather than improve itself. The premise is that the better angels of industry will voluntarily spew less pollution because it's in their financial interest.

The problem is that there's no guarantee the power plant closest to you would get cleaner. Measured nationally, air might be purer; locally, it may not be. The Northeast, downwind of so many dirty power plants, would be especially vulnerable.

New Jersey scrapped an experimental, open-market "cap-and-trade" program last year (after fining a participant $2 million and ordering $3 million in pollution controls). State environmental commissioner Bradley Campbell called it "deeply flawed."

In theory, a national cap-and-trade program could work for some pollutants if it had precisely drawn regions, low caps, and tight monitoring and verification. Before starting, all power plants would have to reach a reasonably clean baseline. Clear Skies lacks those safeguards.

Emissions trading actually works well with carbon dioxide, a greenhouse gas that contributes to global warming. But even though power plants are a top greenhouse polluter, the President is ignoring climate change.

Trading won't work with poisons like mercury, which pool at dangerous levels near their source. Only local cleanup eases health risks.

A report called "America's Children and the Environment," whose findings don't support the Clear Skies time line on mercury, has been collecting dust in an EPA drawer for several months. The EPA finally released it last week after a news leak.

The science is clear: Mercury contamination endangers children. Congress shouldn't be complicit in that risk.

Health & Fitness

VF D5

TUESDAY, JULY 29, 2003

The New York Times

Does Mercury Matter? Experts Debate the Big Fish Question

By JAMES GORMAN

What to eat?

Sometimes it seems to be an impossible question. There are several different kinds of fats to worry about, complex and simple carbohydrates, vitamins in and out of the bottle, chickens that run free before they die and those that don't, grass or grain-fed beef.

At least all the experts agree that fish is good for you. It's high in protein, low in fat, with those terrific omega-3 fatty acids. But then, there's mercury.

Everywhere on the planet, fish are accumulating mercury in their tissues, often as the result of airborne mercury that finds its way into rivers and seas. And mercury, in all its forms, is highly toxic. In fish, it occurs in the form of methylmercury, which is known to damage neurons, particularly developing neurons. The damage seen in humans and animals at high doses is severe. Many studies — though not all — have concluded that low levels can have subtle negative effects as well if certain fish are a major part of the diet.

The Food and Drug Administration and the Environmental Protection Agency have advised that groups most sensitive to methylmercury — women of child-bearing age and young children — not eat swordfish, king mackerel, tilefish and shark, all of them having relatively high mercury levels. Neither agency defines "young," perhaps because the brain continues to develop until early adulthood, but with each year the potential hazard of mercury is less.

Furthermore, the F.D.A., which oversees fish sold in stores and restaurants, advises these women and children that 12 ounces a week of various kinds of fish, is safe. The E.P.A., which is concerned with freshwater fish of the sort anglers bring home, advises the same group to eat no more than six ounces week for adults, or two ounces a week for children.

Researchers started to worry seriously about mercury exposure after several incidents of poisoning in Japan and Iraq. In 1972, about 6,500 people in Iraq were poisoned by methylmercury, which was used in an antifungal treatment for grain seeds. More than 400 died. Some of the treated seeds had been mistakenly turned into flour, and then bread.

The need to find out what low doses of mercury were doing, particularly to developing children, fueled several major studies in the Faeroe Islands in the North Atlantic, the Seychelles in the western Indian Ocean, and New Zealand. Dr. Roberta F. White, chairwoman of the department of environmental health at Boston University and director of the Boston Environmental Hazards Research Center, has been pursuing the Faeroe Islands study for years. She said children exposed to mercury before birth showed clear effects. "The greater the mercury exposure, the poorer they did" on tests that measured nerve functioning, she said.

The effects were subtle, Dr. White said, not something that would be noticed outside a research project, nothing that would be brought to a doctor's attention. She compared the results to some of the damage cause by low-level lead exposure. Similar negative effects were found in New Zealand and in some smaller studies.

Because of these results, she said, "We needed to reconsider how much exposure to methylmercury pregnant women should have."

And exposures were reconsidered. In June, the World Health Organization changed the level of mercury consumption considered safe from 3.6 to 1.5 micrograms per kilogram of body weight per day. (A microgram is 0.000000035 of an ounce, and a kilogram is 2.2 pounds.)

The Environmental Protection Agency has set its reference dose level, the exposure considered to be safe, at 0.1 microgram per kilogram of body weight per day. A National Academy of Sciences report, basing its findings on the Faeroe Islands study, confirmed that this level was safe.

The Food and Drug Administration uses a different measurement, and it has come under considerable criticism from activist groups like the Mercury Policy Project. Its reference dose is 0.4 micrograms per kilogram of body weight.

However, said Dr. David Acheson, the F.D.A.'s chief medical officer concerned with the mercury issue, the number is misleading. It was based on data from Iraq, he said, and is out of date for groups at particular risk. In practice, he said, the important number is the 12 ounces of various fish each week. That advisory was issued in 2001 and is still in effect. The advisory, he said, "is designed to keep at-risk individuals below" the E.P.A.'s recommended dose.

Canned tuna, that lunchbox staple, comes under closer scrutiny.

There are also numerous state health advisories. Melanie Miller, director of communications and marketing for the U.S. Tuna Foundation, an industry group, said: "It's very frustrating for us as an industry. Nobody's telling the same story."

Dr. Acheson said the F.D.A. was working out a joint advisory with the E.P.A. that should be issued by January.

The F.D.A. is also looking into another matter of concern, canned tuna.

Michael Bender, director of the Mercury Policy Project of the Tides Center in Montpelier, Vt., has been a longtime critic of the F.D.A. policies on mercury, and most recently his group's efforts have focused on canned tuna. This is by far the most consumed fish, and he argues that consumers deserve to be warned specifically about its mercury levels.

His group tested 48 cans of white or albacore tuna and found mercury levels averaging 0.5 parts per million, significantly higher than F.D.A. data. The F.D.A. and the Tuna Foundation agree that albacore has more mercury than light tuna, which comes from smaller fish, but set the figure at [illegible] parts per million compared with 0.13 for light tuna, Dr. Acheson said.

Mr. Bender said consumers, particularly women and children, should be warned separately about tuna, since it is so popular. At 0.5 p.p.m. he said, a woman weighing 132 pounds who ate only 6 ounces of tuna in a week would be getting 1.4 micrograms per kilogram of body weight in that week — too much mercury by E.P.A. standards.

Dr. Acheson said that the F.D.A. was starting more intense testing of tuna and that it would consider whether to say anything specific about canned tuna based on those findings.

This week the agency is conducting meetings of so-called stakeholders — including

Continued on P[illegible]

Experts Debate the Threat Posed by Mercury in Fish

Continued From Page 5

the Tuna Foundation and the Mercury Policy Project — to discuss F.D.A. policy and advisories on mercury in fish.

Ms. Miller of the Tuna Foundation said that her group thought no new warnings were necessary and that the current F.D.A. advisory adequately protected consumers.

Despite the warnings, the results of one study present a challenge to the notion that low levels of mercury have any negative effect. Dr. Gary J. Myers, a professor of neurology and pediatrics at the University of Rochester Medical Center, followed children in the Seychelles whose mothers ate about 12 meals a week of fish with mercury levels comparable to those in fish eaten by consumers in the United States.

Dr. Myers and other researchers did tests that were sensitive enough to pick up the effects on neurological performance of things like poverty and parents' educational level. They found no adverse effect from prenatal mercury exposure.

Dr. Myers, who began studying mercury in Iraq, said that in the course of nine years of study in the Seychelles, "We've really not been able to confirm that there are adverse effects" at consumpton levels far above the average American's.

The differences between this study and the Faeroe Islands study have puzzled other scientists. Dr. White, of Boston University, suggested that measures of mercury in umbilical cord blood correlated better with neurological effects than measures in hair. Dr. Myers said there was no scientific evidence to suggest that the blood measures were better than the hair measures.

Dr. Myers also pointed out that the Seychelles population ate fish, whereas the Faeroes population got a substantial amount of its mercury exposure from whale meat. Few people in the United States eat whale meat, he noted.

The National Academy of Sciences panel considered both studies, and others. Dr. Alan Stern, chief of the bureau for risk analysis in the New Jersey Department of Environmental Protection, was on the panel. He said the findings in the Seychelles and the Faeroes were based on sound research. The panel recommended using the Faeroes study to set acceptable levels of mercury intake because other studies had also found neurological problems, and the panel did not want to use one study that found no effect as a benchmark.

This kind of seeming discrepancy occurs all the time in epidemiology, he said. Some research uncovers deleterious effects of chemicals under scrutiny, and other research does not. "From a public health standpoint," Dr. Stern said, if there are well-done studies showing a danger to the public, the best approach is to be cautious and to "make prudent recommendations based on those studies."

Tip the Scale in Favor of Fish: The Healthful Benefits Await

By JANE E. BRODY

The word fish has entered the American vocabulary in several not so savory ways. "That's a fish tale," for example, and "That smells fishy," suggest a highly exaggerated story or blatant lie or misdeed.

But in the last two decades, fish and shellfish have enjoyed a more positive press, lauded for conveying a number of important health benefits to those who regularly dine on them.

Indeed, the new image of fish has sparked all manner of creativity, with fish and shellfish starring on platters, napkin rings, glassware, jewelry, towels, clothing and even carpets.

Sadly, this popularity of fish as an artistic form has not translated into any drastic transformation in American dining habits. Seafood hardly holds its own next to red meat and poultry as a source of animal protein, and the industry says consumption peaked in 1987 at 16.2 pounds, falling just below 14.8 pounds in 2001.

I say sadly, because evidence suggests that eating just two fish meals a week — two three-ounce servings — can make a significant difference in the risk of developing ills like heart attacks and strokes.

Part of the benefit may come simply from eating fish in place of red meat, a supposed culprit in cardiovascular disease. But another more important part, the evidence suggests, involves the fats found in fish, namely omega-3 fatty acids, which may have a physiological role in preventing cardiovascular disease.

But as with most worthwhile things in life, with fish and shellfish, there's no free lunch. Risks as well as benefits attend eating seafood, although a wise and well-informed consumer can certainly tip the balance well in favor of the benefits.

Fish Oils and Health

For fish to maintain fluidity in cold water, their fats have to remain liquid, and liquid fats (really oils) are polyunsaturated. But fish oils, rich in omega-3 fatty acids, are chemically different from the polyunsaturated oils in plants like corn and soybeans, and it is that difference that has given fish star billing.

The two omega-3's in fish are eicosapentaenoic acid, or EPA, and docosahexaenoic acid, or DHA. They are considered essential fatty acids, although they can be formed in the body from another omega-3, alpha-linolenic acid, or ALA. It is found in plants like flaxseed, spinach, mustard greens, soybeans, canola oil, wheat germ and walnuts, as well as in marine animals that eat plants containing ALA.

The conversion rate is poor, however, and you will have to consume a large amounts of ALA to obtain a meaningful amount of EPA and DHA. Eating fish is far more efficient.

DHA is a natural ingredient in breast milk, and it is critical to the normal development of the brain and retina. It has recently been approved as an additive to infant formula. In addition, the omega-3 acids perform many biochemical functions that can benefit the heart and blood vessels.

They can inhibit the synthesis of substances that promote inflammation, reduce the tendency of the blood to form clots, stabilize the electrical activity of the heart, lower triglyceride levels, reduce blood pressure moderately and improve the functioning of artery linings.

Most, but not all, studies have found that people who eat fish regularly experience significant reductions in the risk of heart attacks, strokes and deaths from all causes. The benefit has been most clear-cut among people who already have cardiovascular disease, but it has also been found among those who are initially healthy.

In one study, of 334 people who had suffered first heart attacks and 493 healthy people matched for age, sex and location, eating just 5.5 grams of omega-3 fatty acids a month was associated with a 50 percent reduction in the risk of cardiac arrest.

That amount of omega-3's can come from just four three-ounce servings of Atlantic salmon a month. The omega-3's have proved especially beneficial for people with Type 2 diabetes, who have a greatly increased risk of developing heart disease.

Other suggested benefits include an anti-inflammatory effect that can help people with autoimmune diseases like rheumatoid arthritis, psoriasis and ulcerative colitis.

The various findings prompted the American Heart Association to recommend in October 2000 that everyone strive to consume at least two fish servings a week, especially fatty fish like salmon, sardines, mackerel, herring, lake trout, tuna and anchovies.

Also Consider the Risks

Some fish are contaminated with mercury and other toxic substances introduced as industrial pollutants into their waters. An international research team reporting in The New England Journal of Medicine in November noted that the mercury content of some fish might diminish their cardioprotective effects. Again, not every study has shown this.

Still, mercury is a well-known neurologic and kidney toxin that is best not ingested. The fish most likely to be contaminated are large deep-sea species like swordfish, king mackerel, shark and tilefish and are best avoided, especially by pregnant women.

Local health departments periodically issue advisories about contaminated species. Mercury at levels experienced by Americans may also cause problems with fine motor control and memory in adults, a new study by Dr. Edna M. Yokoo and colleagues at Johns Hopkins says.

Fish, especially shellfish, can become contaminated by harmful micro-organisms that occur naturally and those that result from sewage pollution. A natural toxin produced by marine organisms called dinoflagellates can cause ciguatera poisoning, a potential hazard from large fin fish, especially barracuda that frequent tropical reefs.

But it is raw fish, especially raw shellfish, that is the most common source of food poisoning. Only ocean-dwelling fish should be used in sushi, sashimi and ceviche, and only from reliable sources that know how to spot contamination by parasites.

The filter-feeders — clams, oysters and other mollusks like mussels and scallops — can accumulate waterborne bacteria, viruses or toxins that can cause severe gastrointestinal problems when the shellfish is eaten uncooked.

Crustaceans like shrimp and lobster should never be eaten raw, and all mollusks that are eaten should have been harvested from certified beds. The safest bet, however, is to cook them. For example, oysters and clams should be boiled for at least three minutes or steamed for four to six minutes to destroy harmful micro-organisms.

Also important is to handle fish and shellfish safely in the home, just as you should all poultry and red meat. Raw seafood should be kept cold, at refrigerator temperatures, even while defrosting, and care must be taken to keep the juices of raw seafood from contaminating other foods, especially those eaten raw.

Furthermore, all utensils and surfaces used to prepare raw seafood should be thoroughly washed with hot soapy water before being used for any other purpose. The tray or platter that held the raw seafood also should be washed before it's used for the cooked seafood.

Finally, there is the problem of allergies. About 1 percent of adults have clinically proven food allergies, and shellfish are among the most common culprits. Those with known allergies to fish or shellfish must take particular care to avoid all sources and would be wise to carry a loaded syringe of Adrenalin like an EpiPen at all times to counter life-threatening reactions to inadvertent exposure. Even kissing someone who just ate shrimp, for example, can cause a reaction in someone allergic to it.

Fish in a Capsule Can Fill Nutrition Gap

For people who do not eat enough fish and shellfish to consume helpful amounts of fish oils, countless supplements can fill the gap.

For people with established heart disease and those known to face a high risk of developing it, the recommended consumption of about one gram a day of the omega-3 fatty acids in fish oils may prove challenging without the supplements.

The July issue of Consumer Reports describes its evaluation of 16 products that were laboratory-tested for both omega-3 content and contaminants. Because the magazine found "no significant differences in the quality or purity of these supplements," it recommends selecting a fish-oil supplement based on price.

The cheapest come from two large warehouse companies, the Kirkland Signature Natural Fish Oil from Costco and Member's Mark Omega 3 Fish Oil from Sam's Club. The next cheapest brand, Spring Valley Natural Fish Oil Concentrate, costs three times as much per daily dose of one gram a day of EPA and DHA combined.

But as with seafood itself, fish oils can also cause problems. Some people with bleeding disorders and those who take anticoagulants, including aspirin, vitamin E, garlic and ginkgo as well as warfarin (Coumadin), may increase their risk of bleeding by taking fish-oil supplements.

To reduce the problem of a fishy odor from supplements, the concentrated oils are encapsulated. But that does not protect against a fishy odor when burping. Thus, the capsules are best taken at night or at the start of a meal, but not on an empty stomach.

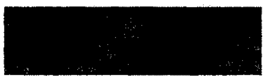
A Healthful, Sustainable Meal (Hold the Mercury)

BENEFICIAL OILS Studies have confirmed the benefits of certain omega-3 oils, found only in fish, in protecting against cardio-vascular disease. One study found that 5.5 grams a month reduced the risk of heart attack by 50 percent.

These fish are ranked by the amount of omega-3 oils in a cooked 4-ounce serving.

2 OR MORE GRAMS OF OMEGA-3S

Pacific herring	2.4 g
Atlantic herring	2.3
Pacific or jack mackerel	2.1
Atlantic salmon	2.1
Sablefish	2.0



1-2 GRAMS

Canned pink salmon	1.9 g
Whitefish	1.9
Pacific oysters	1.6
Pink salmon	1.5
Atlantic mackerel	1.4
Sockeye or red salmon	1.4
Coho salmon	1.2
Bluefish	1.1
Trout	1.1
Eastern oysters	1.0
Rainbow smelt	1.0
Whiting (hake)	1.0

LESS THAN 1 GRAM

Freshwater bass	0.9 g
Blue mussels	0.9
Swordfish	0.9
Rainbow trout	0.8
White canned tuna	0.8
Canned sardines	0.7
Flounder or sole	0.6
Halibut	0.5
Rockfish	0.5
Shrimp	0.4
Snapper	0.4
Sturgeon	0.4
Atlantic perch	0.3
Clams, fresh or canned	0.3
Haddock	0.3
Light canned tuna	0.3
Yellowfin tuna	0.3
Atlantic cod	0.2
Catfish	0.1

LOW OR NO OILS Imitation crab, made from low-fat pollock, seems to lose its omega-3 during processing, as do fish sticks and fast-food fish sandwiches.

A SUSTAINABLE FISH CHECKLIST The Monterey Bay Aquarium in California has compiled a guide to fish it recommends and those it advises comsumers to avoid, based on the following:

BYCATCH If there is significant waste of other species in the fishing process.
DEPLETION Whether the species has been overfished.
FARMING Whether safe, sustainable farming techniques are used.
HABITAT Whether habitats are damaged or destroyed in the fishing process.

BEST CHOICES	CAUTION	AVOID
Abalone (farmed)	Clams (wild-caught)	Caviar (Beluga/Osetra/Sevruga)
Catfish (U.S. farmed)	Cod, Pacific	Chilean sea bass
Caviar (farmed)	Crab, imitation/surimi	Cod, Atlantic/Icelandic
Clams (farmed)	Crab, king (Alaska)	Crab, king (imported)
Crab, Dungeness	Crab, snow	Lingcod
Halibut, Pacific	Lobster, American	Monkfish
Hoki	Mahi-mahi	Orange roughy
Lobster, rock (Calif., Australia)	Mussels (wild-caught)	Rockfish (rock cod/Pacific snapper)
Mussels (farmed)	Oysters (wild-caught)	Salmon (farmed/Atlantic)
Oysters (farmed)	Pollock	Sharks (except West Coast thresher)
Sablefish/black cod (Br. Columbia, Alaska)	Sablefish/black cod (Calif. or Wash.)	Shrimp (imported)
Salmon, canned	Salmon (wild-caught, Ore. or Wash.)	Sturgeon, wild-caught
Salmon (wild-caught, Calif. and Alaska)	Sand dabs	Swordfish, Atlantic
Sardines	Scallops, bay (farmed)	Tuna, bluefin
Sea bass, white	Scallops, sea	
Shrimp/prawns (trap-caught)	Shark, thresher (U.S. West Coast)	
Squid (Calif.)	Shrimp (U.S. farmed or wild)	
Striped bass (farmed)	Sole, English/petrale/Dover/rex	
Sturgeon (farmed)	Swordfish, (U.S. W. Coast)	
Tilapia (farmed)	Tuna, canned	
Trout, rainbow (farmed)	Tuna, albacore/yellowfin/bigeye (caught by long line or purse seine)	
Tuna, albacore/yellowfin/bigeye (troll- or pole-caught)		

Shen Q...

CONTAMINANTS There is widespread disagreement about ... mercury and how much people can tolerate. Mercury is measured in parts per million but Food and Drug Administration advisories for safe consumption dep... weight and other factors. Here is the F.D.A. ... fish ...

The F.D.A. says young children and women of childbearing age should avoid these four fish, and limit consumption to 12 oz. of all fish per week.

Tilefish	ppm 1.45
Swordfish	1.00
Shark	0.96
King mackerel	0.73
Grou... (Mycteroperca)	0.43
Tun...	0.32
Northern ... (U.S.)	0.31
Grouper ...	0.27
Halibut	0.23
Sablefish	0.22
Pollock	0.20
Dungeness crab	0.18
Tuna ...	0.17
Blue...	0.17

Associated Press

Sources: Center for Scie... the Public Interest; U.S. Department of Agriculture; ...

August 18, 2003.

Hon. JAMES M. INHOFE, *Chairman,*
Committee on Environment and Public Works,
U.S. Senate,
Washington, DC.

DEAR MR. CHAIRMAN: Thank you for offering me the opportunity to respond to certain comments that were made in the EPW committee hearing on Tuesday, July 29, of this year. I hope I can clear up any confusion that might have been caused by incomplete, misleading or erroneous testimony that day.

The testimony in question by Dr. Michael Mann stated:

> "It's unfortunate to hear comments about the supposed inconsistencies of the satellite record voiced here years after that has been pretty much debunked in the peer-reviewed literature in Nature and Science. Both journals have, in recent years, published . . . articles indicating that in fact, the original statement that the satellite record showed cooling was flawed because . . . the original author, John Christy, did not take into account a drift in the orbit of that satellite, which actually leads to a bias in the temperatures . . . Christy and colleagues have claimed to have gone back and fixed that problem. But just about every scientist who has looked at it says that this fix isn't correct and that if you fix it correctly then the satellite record actually agrees with the surface record, indicating fairly dramatic rates of warming in the past two decades."

Virtually all of this testimony is misleading or incorrect. I will touch on the major problems, point-by-point, and I will try to be brief.

1. Certainly no one has "debunked" the accuracy of the global climate dataset that we built at The University of Alabama in Huntsville (UAH) using readings taken by microwave sensors aboard NOAA satellites. This dataset has been thoroughly and rigorously evaluated, and has been published in a series of peer-reviewed papers beginning in *Science* (March 1990). The most recent version of the dataset was published in May 2003 in the *Journal of Atmospheric and Oceanic Technology* after undergoing a strenuous peer review process.

2. We, and others, are constantly scrutinizing our techniques to find ways to better analyze the data. In every case except one we discovered needed improvements ourselves, developed a method for correcting the error, and published both the error and the correction in peer-reviewed journals. When Wentz, *et al.* (1998) published their research on the effects of orbital decay (the one exception) they explained an effect we immediately recognized, but which was partially counterbalanced by other factors we ourselves discovered. Since that time we have applied the corrections for both orbital decay and other factors, and have published the corrected data in peer-reviewed journals.

3. The UAH satellite record does not show cooling in the lower troposphere and hasn't shown a long-term cooling trend since the period ending in January 1998. I cannot say where this chronic cooling misconception originated. Our long-term data show a relatively modest warming in the troposphere at the rate of 0.133° Fahrenheit per decade (or 1.33° Fahrenheit per century) for the period of November 1978 to July 2003.

4. There is no credible version of the satellite dataset that "actually agrees" with the surface temperature record for the past 25 years, nor one that shows "fairly dramatic rates of warming." The as-yet-unexplained differences between the surface and satellite data are at the heart of the controversy over the accuracy of the satellite data.

While much of the surface data remains uncalibrated and uncorroborated, we have evaluated our UAH satellite data against independent, globally-distributed atmospheric data from the U.S. and the U.K. (Hadley Centre) as shown in the figure (enclosure 1). We published the results of those comparisons in numerous peer-reviewed studies (enclosure 2). In each case we found excellent consistency between the satellite data and the atmospheric data. One should note that such independent corroboration has not been performed on the other satellite temperature datasets alluded to in the quoted testimony.

This consistency between two independent datasets gathered using very different techniques gives us a high level of confidence that the UAH satellite dataset provides a reliable measure of global atmospheric temperatures over more than 90 percent of the globe. (By comparison, one of the most often quoted surface temperature datasets achieves partial-global coverage only by claiming that certain isolated thermometer sites provide representative temperatures for an area roughly equaling two-thirds of the contiguous 48 states, an area that would reach from about Brownsville, Texas, to Grand Forks, North Dakota.)

5. A final point relates to numerous comments elsewhere in the testimony in which an appeal to a nebulous "mainstream climate community" was made to support what was stated. First, the notion that "thousands" of climate scientists agreed on the IPCC 2001 text is an illusion. I was a lead author of IPCC 2001, as was Dr. Mann. There were 841 lead authors and contributors, the majority of whom were not climatologists and who provided input in the area in which they have expertise only to their tiny portion of the 800+ page document. These 841 were not asked to approve nor where they given the opportunity to give a stamp of approval on what was finally published.

Although I might be outside the "mainstream," according to Dr. Mann's perspective, I have never thought a scientist's goal was to achieve membership in the "mainstream." My goal is to produce the most reliable climate datasets for use in scientific research. Whether they show warming or cooling is less important to me than their reliability and accuracy. That these datasets have been published in numerous peer-reviewed venues is testimony to accomplishing this goal and, by inference, would place me inside the mainstream climate community. In addition to being an IPCC lead author, significant achievement awards from NASA and the American Meteorological Society along with my recent election as a Fellow of the AMS are evidence of my impact on the community of scientists.

I hope this clears up any confusion you or your committee members might have had about the UAH global temperature data. If you or any of your committee members have any questions, I will be delighted to answer them to the best of my ability.

Thank you again for offering me this opportunity. I remain,

Sincerely,

JOHN CHRISTY, PH.D.

○